

INTERNATIONAL HANDBOOK SERIES ON ENTREPRENEURSHIP

---

# The Life Cycle of Entrepreneurial Ventures

*Edited by*  
Simon C. Parker

---

 Springer

THE LIFE CYCLE  
OF ENTREPRENEURIAL VENTURES

# International Handbook Series on Entrepreneurship

---

## VOLUME 3

---

### *Series Editors*

Zoltan J. Acs, *University of Baltimore, U.S.A.*

David B. Audretsch, *Indiana University, U.S.A.*

### SERIES FOREWORD

Interest in entrepreneurship has surged in the last decade. Scholars across a broad spectrum of fields and disciplines have responded by generating new research approaches uncovering a wealth of new findings and insights about entrepreneurship. This new research spans not just a diverse set of fields, such as management, finance, psychology, economics, sociology, and geography but also a wide range of countries reflecting the fact that entrepreneurship is a global phenomenon. The exceptionally cross-disciplinary nature of entrepreneurship has made it difficult for scholars in any one particular field to become aware of and understand the leading contributions and insights emerging in other disciplines. The purpose of this series is to compile a series of handbooks, each devoted to particular issue in the entrepreneurship field. Each handbook will draw upon the leading international scholars from the entire range of disciplines contributing to entrepreneurship to articulate the state of knowledge about a particular topic. The contribution should identify the fundamental questions which are being posed, the methodological approaches, types of data bases used for empirical analyses, the most important empirical regularities to emerge in the literature, major policy conclusions, and the most promising research directions. Thus, each handbook will reflect the interdisciplinary nature of entrepreneurship that has proven to be elusive to discipline-based scholars. A goal of the Handbook Series is not only to provide a state-of-the-art coverage of what has been learned about entrepreneurship, but that when viewed in its entirety, entrepreneurship is emerging as a bona fide academic discipline.

The particular topics in the Series will be drawn from discussions with the leading scholars. Each handbook will be directed and compiled by a Handbook Editor. (S)he will work closely with the Series Editor to ensure that the contents and contributions are appropriate, and that there is consistency with the other volumes in the Series.

The titles published in this series are listed at the end of this volume.

# THE LIFE CYCLE OF ENTREPRENEURIAL VENTURES

*Edited by*

SIMON PARKER  
*Durham Business School*

 Springer

Library of Congress Control Number 2006920279

ISBN-10: 0-387-32156-X  
ISBN-13: 978-0387-32156-1

e-ISBN-10: 0-387-32313-9  
e-ISBN-13: 978-0387-32313-8

Printed on acid-free paper.

© 2006 Springer Science+Business Media, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

springer.com

# Table of Contents

|  |      |
|--|------|
| List of Contributors .....   | ix   |
| List of Figures .....  | xiii |
| List of Tables .....   | xv   |
| <b>1. Introduction</b>   |      |
| Simon C. Parker .....  | 1    |
| <b>Stage 1: Beginnings</b>   |      |
| <b>2. Nascent Entrepreneurs</b>  |      |
| Joachim Wagner .....   | 15   |
| <b>3. New Venture Strategies: Transforming Caterpillars into Butterflies</b> |      |
| Shaker A. Zahra .....  | 39   |
| <b>Stage 2: Aspects of Entry and New Venture Creation</b>                    |      |
| <b>4. Entrepreneurship Education—A Compendium of Related Issues</b>          |      |
| Lena Lee and Poh-Kam Wong .....  | 79   |
| <b>5. Immigration, Entrepreneurship and the Venture Start-up Process</b>     |      |
| Herbert J. Schuetze and Heather Antecol .....                                | 107  |

|   |     |
|---|-----|
| <b>6. Location and New Venture Creation</b>   |     |
| David B. Audretsch and Erik Lehmann .....   | 137 |
| <b>7. On Factors Promoting and Hindering Entry and Exit</b>                                     |     |
| Martin A. Carree .....  | 161 |
| <b>Stage 3: Financing Ventures</b>  |     |
| <b>8. Debt Finance and Credit Constraints on SMEs</b>   |     |
| Robert Cressy .....   | 187 |
| <b>9. Public Policy, Start-up Entrepreneurship and the Market for Venture Capital</b>           |     |
| Christian Keuschnigg and Søren Bo Nielsen .....   | 227 |
| <b>10. Informal Sources of Venture Finance</b>  |     |
| Colin M. Mason .....  | 259 |
| <b>11. Microfinance and Poor Entrepreneurs</b>  |     |
| João Pedro Azevedo .....  | 301 |
| <b>Stage 4: Venture Development I: Private Sector Issues</b>                                    |     |
| <b>12. Entrepreneurs as Producers</b>   |     |
| Simon C. Parker .....   | 337 |
| <b>13. What Do We Know About Small Firm Growth?</b>   |     |
| Per Davidsson, Leona Achtenhagen and Lucia Naldi .....  | 361 |
| <b>Stage 5: Venture Development II: Social Issues</b>   |     |
| <b>14. Nonprofit Social Entrepreneurship</b>  |     |
| Helen Haugh .....   | 401 |
| <b>15. Entrepreneurship among Disadvantaged Groups: Women, Minorities and the Less Educated</b> |     |
| Robert W. Fairlie .....   | 437 |
| <b>Stage 6: Venture Performance and Harvesting</b>  |     |
| <b>16. Early Stage Survival and Growth</b>  |     |
| Marc Cowling .....  | 479 |

|  |     |
|--|-----|
| <b>17. Venture Performance and Venture Inputs: The Role of Human and Financial Capital</b> |     |
| C. Mirjam van Praag .....  | 507 |
| <b>18. Harvesting in High Growth Firms</b>   |     |
| Maks Tajnikar, Petra Došenovič Bonča and Lidija Zajec .....                                | 533 |
| <b>Index</b> .....   | 583 |



## List of Contributors

LEONA ACHTENHAGEN  
Jönköping International Business School  
Sweden  
E-mail: leona.achtenhagen@jibs.hj.se

HEATHER ANTECOL  
Department of Economics  
Claremont McKenna College  
Claremont, California  
USA  
E-mail: heather.antecol@  
claremontmckenna.edu

DAVID B. AUDRETSCH  
Indiana University and Max Planck  
Institute of Economics,  
Entrepreneurship, Growth and  
Public Policy  
Kahlaische Straße 10  
D-07745 Jena  
Germany  
E-mail: audretsch@econ.mpg.de

JOÃO PEDRO AZEVEDO  
Brazilian Institute of Applied Economic  
Research  
Av. Pres. Antonio Carlos, 51-10° andar  
Rio de Janeiro-RJ, 20020-010  
Brazil  
E-mail: jazevedo@ipea.gov.br

PETRA DOŠENOVIČ BONČA  
Faculty of Economics, University of  
Ljubljana  
Kardeljeva ploščad 17  
1109 Ljubljana

Slovenia  
E-mail: petra.d.bonca@ef.uni-lj.si

MARTIN A. CARREE  
Department of Organization and Strategy  
University of Maastricht  
FdEWB  
PO Box 616  
6200 MD Maastricht  
The Netherlands  
E-mail: m.carree@os.unimaas.nl

MARC COWLING  
Principal Economist  
Institute for Employment Studies  
Mantell Building  
University of Sussex  
Falmer  
Brighton BN1 9RF  
United Kingdom  
E-mail: marc.cowling@ies.ac.uk

ROBERT CRESSY  
Cass Business School  
London EC1Y 8TZ  
United Kingdom  
E-mail: r.c.cressy@city.ac.uk

PER DAVIDSSON  
Brisbane Graduate School of Business  
Queensland University of Technology  
(and Jönköping International  
Business School)  
Gardens Point, Brisbane  
4001 Queensland  
Australia  
E-mail: per.davidsson@qut.edu.au

ROBERT W. FAIRLIE  
Department of Economics  
University of California  
Santa Cruz, CA 95060  
USA  
E-mail: rfairlie@ucsc.edu

HELEN HAUGH  
The Judge Institute of Management  
University of Cambridge  
Trumpington Street  
Cambridge CB2 1AG  
United Kingdom  
E-mail: h.haugh@jims.cam.ac.uk

CHRISTIAN KEUSCHNIGG  
University of St Gallen  
Varnbuelstrasse 19  
CH-9000 St. Gallen  
Switzerland  
E-mail: christian.keuschnigg@unisg.ch

LENA LEE  
National University of Singapore  
Entrepreneurship Centre  
E3A Level 6  
10 Kent Ridge Crescent  
Singapore 119260  
E-mail: lenalee@nus.edu.sg

ERIK LEHMANN  
University of Augsburg  
Chair of Management and Organisation  
Universitätsstraße 16  
D-86159 Augsburg  
Germany  
E-mail: Erik.Lehmann@wiwi.uni-  
augsburg.de

COLIN M. MASON  
Hunter Centre for Entrepreneurship  
University of Strathclyde  
Glasgow G1 1XH  
United Kingdom  
E-mail: colin.mason@strath.ac.uk

LUCIA NALDI  
Jönköping International Business School  
Sweden  
E-mail: lucia.naldi@jibs.hj.se

SØREN BO NIELSEN  
Copenhagen Business School  
Solbjerg Plads 3  
DK-2000 Frederiksberg  
Denmark  
E-mail: sbn.eco@cbs.dk

SIMON C. PARKER  
Director of the Centre for  
Entrepreneurship  
Durham Business School  
Mill Hill Lane  
Durham DH1 3LB  
UK  
E-mail: s.c.parker@durham.ac.uk

C. MIRJAM VAN PRAAG  
University of Amsterdam and Tinbergen  
Institute  
Roetersstraat 11  
1018WB Amsterdam  
The Netherlands  
E-mail: c.m.vanpraag@uva.nl

HERBERT J. SCHUETZE  
Department of Economics  
University of Victoria  
Victoria, BC  
Canada  
E-mail: hschuetz@uvic.ca

MAKS TAJNIKAR  
Faculty of Economics, University of  
Ljubljana  
Kardeljeva ploščad 17  
1109 Ljubljana  
Slovenia  
E-mail: maks.tajnikar@ef.uni-lj.si

JOACHIM WAGNER  
Institute of Economics  
University of Lueneburg  
Campus 4.210  
D-21332 Lueneburg  
Germany  
E-mail: Wagner@uni-lueneburg.de

POH KAM WONG  
14 Prince George's Park  
NUS Entrepreneurship Centre  
National University of Singapore  
118412 Singapore

SHAKER A. ZAHRA  
Center for Entrepreneurial Studies &  
Department of Strategic  
Management & Organization

Room 3-428, Carlson School of  
Management  
University of Minnesota  
321 19th Ave. South  
Minneapolis, MN 55455  
USA  
E-mail: szahra@csom.umn.edu

LIDIJA ZAJEC  
Faculty of Economics, University of  
Ljubljana  
Kardeljeva ploščad 17  
1109 Ljubljana  
Slovenia  
E-mail: lidija.zajec@ef.uni-lj.si

## List of Figures

|     |   |     |
|-----|---|-----|
| 3-1 | Steps followed in mapping out new venture strategy<br>.....                                 | 42  |
| 7-1 | The life cycle of an industry: five stages<br>.....   | 169 |
| 7-2 | The Dixit (1982) game tree<br>.....   | 176 |
| 8-1 | The frequency and importance of finance sources<br>.....                                    | 191 |
| 8-2 | A motive for the overdraft: timing of the firm's cash flows<br>.....                        | 194 |
| 8-3 | Assets, entrepreneurial ability and credit constraints à la Evans<br>and Jovanovic<br>..... | 207 |
| 8-4 | Effects of control aversion on the amount of borrowing<br>.....                             | 214 |
| 9-1 | Events and notation<br>.....  | 234 |
| 9-2 | Effort and advice<br>.....  | 239 |

|      |   |     |
|------|---|-----|
| 9-3  | The start-up decision                             | 241 |
| 9-4  | Equilibrium entrepreneurship                      | 242 |
| 11-1 | Risk, cost of capital and market failure          | 306 |
| 12-1 | Labor supply curves                               | 352 |
| 13-1 | Alternative foci for studies of small firm growth | 386 |

## List of Tables

|     |   |     |
|-----|---|-----|
| 2-1 | Share of nascent entrepreneurs in the adult population<br>(18–64 years) in 2003<br>.....    | 18  |
| 2-2 | Impact of selected factors on nascent entrepreneurship<br>(29 GEM countries, 2001)<br>..... | 22  |
| 5-1 | Self-employment rates of men<br>.....   | 117 |
| 5-2 | Summary of selected trends in predicted values<br>.....                                     | 119 |
| 5-3 | Decomposition of changes in the probability of self-employment<br>.....                     | 120 |
| 5-4 | Average log weekly earnings of self-employed men<br>.....                                   | 125 |
| 5-5 | Earnings regressions—assimilation, cohort and period effects<br>.....                       | 126 |
| 5-6 | Variable definitions<br>.....   | 130 |
| 6-1 | Descriptive statistics<br>.....   | 150 |

|      |   |     |
|------|---|-----|
| 6-2  | Correlation matrix<br>.....   | 151 |
| 6-3  | Negative binomial and ordered probit regressions estimating<br>geographic proximity to an university<br>..... | 152 |
| 6-4  | Negative binomial regressions estimating geographic proximity<br>to an university<br>.....                    | 153 |
| 6-5  | Results from the semi-parametric Cox regression<br>.....  | 154 |
| 7-1  | Total entrepreneurial activity rates (TEA) across countries, 2002<br>.....                                    | 163 |
| 7-2  | Entry and exit rates for U.S. industries (2nd digit NAICS)<br>.....   | 166 |
| 7-3  | Summary of empirical studies into entry and exit determinants<br>.....  | 167 |
| 8-1  | Sources of small firm finance<br>.....  | 196 |
| 9-1  | Effects of taxes and subsidies<br>.....   | 245 |
| 10-1 | Differences between types of business angel: I—Coveney and<br>Moore<br>.....                                  | 265 |
| 10-2 | Differences between types of business angel: II—Sørheim and<br>Landström<br>.....                             | 266 |
| 10-3 | Stages in the business angel's investment decision<br>.....   | 274 |
| 11-1 | Active loans in alternative financial institutions (in thousands)<br>.....                                    | 315 |

|      |  |     |
|------|--|-----|
| 11-2 | Savings accounts in alternative financial institutions<br>(in thousands)<br>.....      | 315 |
| 11-3 | Performance indicators of selected microfinance institutions by<br>region<br>.....     | 320 |
| 13-1 | Total and organic growth for high growth firms of different sizes<br>.....             | 376 |
| 14-1 | Percentage of paid staff and volunteers by field of activity<br>.....                  | 403 |
| 14-2 | Paid staff and volunteers by country<br>.....  | 407 |
| 14-3 | Civil society sector FTE revenue<br>.....  | 414 |
| 15-1 | Self-employment rates by gender, race, immigration and<br>education<br>.....           | 440 |
| 15-2 | Male and female self-employment rates by country<br>(nonagricultural sectors)<br>..... | 441 |
| 15-3 | Self-employment rates by race/ethnicity for selected countries,<br>2000–2001<br>.....  | 443 |
| 15-4 | Transition matrices by gender, race, immigration and education<br>.....                | 446 |
| 15-5 | Logit regressions for probability of entry into self-employment<br>.....               | 450 |
| 15-6 | Logit regressions for probability of exit from self-employment<br>.....                | 457 |
| 15-7 | Decomposition of racial/ethnic gaps in self-employment entry<br>rates<br>.....         | 463 |



|      |  |     |
|------|--|-----|
| 15-8 | Decomposition of racial/ethnic gaps in self-employment exit rates<br>.....     | 465 |
| 16-1 | Review of empirical studies of factors affecting business survival<br>.....    | 486 |
| 16-2 | Review of empirical studies of factors affecting business growth<br>.....      | 494 |
| 17-1 | Evidence about the effects of financial capital on performance<br>.....        | 519 |
| 17-2 | Estimation results: the effects of capital constraints on performance<br>..... | 524 |

SIMON C. PARKER

*Director of the Centre for Entrepreneurship, Durham Business School, Mill Hill Lane, Durham  
DH1 3LB, UK*

*E-mail: s.c.parker@durham.ac.uk*

## 1. Introduction

The theme of this volume is the life cycle of entrepreneurial ventures. The term “life cycle” refers to the sequence of stages in the evolution of new ventures. These stages commence with the intentions and actions of nascent entrepreneurs, and the creation of the venture as a new organization. They are accompanied by the acquisition of necessary financial and nonfinancial resources. As ventures develop, their owners remain instrumental in shaping their performance; and the venture development process culminates in involuntary exit or purposive harvesting by the entrepreneur. These stages furnish a natural framework for organizing the chapters that comprise this volume.

The venture life cycle offers a useful framework within which to analyse entrepreneurship. While it is, of course, not the only standpoint from which to approach the subject, it helps to organize and clarify the varied decisions that entrepreneurs make, the rationale for various organizations and institutions observed around us, and how the conjunction of entrepreneurs and organizations shapes social and economic outcomes.

This volume shares several features with its predecessors in the *Handbook* series. First, the volume takes an avowedly interdisciplinary approach. It contains chapters written from a variety of disciplinary backgrounds, including economics, strategy, finance, industrial organisation, and sociology. Second, the list of contributors and the evidence base from which the contributors have drawn is truly international. No fewer than a dozen different countries are represented among the list of contributors, and every author cites evidence relating to a range of countries. Third, every chapter provides a self-contained overview of the field to help scholars quickly familiarise themselves with the subject at hand. Fourth, all chapters have been anonymously peer reviewed and carefully revised before finally being accepted for publication.

There are also a few differences with previous volumes. First, the sequential life cycle structure of this volume is novel. Second, many of the chapters contain new theoretical or empirical findings in addition to overviews and syntheses. Indeed, some of the chapters combine theory and evidence to better justify the former and motivate the latter. This was not planned as part of a deliberate editorial strategy, but was an outcome that emerged naturally and organically as the volume developed. Third, some interesting and topical issues that currently lack accessible, up-to-date and concise literature reviews are included, namely on nascent entrepreneurship, entrepreneurship education, informal sources of venture capital, social entrepreneurship and harvesting. It is hoped that these features will enhance the appeal and usefulness of the volume.

Of course, one volume cannot hope to cover every topic relating to entrepreneurship and venture life cycles. It is important to note at the outset that the life cycle of new entrepreneurial ventures is related to, but distinct from, the life cycle of products and industries. Evolutionary perspectives regarding product and industry life cycles are discussed in one or two places below but lack a dedicated chapter of their own—reflecting this editor’s preference for a focus on a sequential life cycle structure of ventures specifically rather than industries as a whole. Also, the volume does not dedicate chapters to any of the following topics: the psychology of entrepreneurs, theories of the firm, innovation or entrepreneurship by employees of existing firms (“intrapreneurship”). These issues have been covered in previous volumes of the *Handbook* series; I have tried to avoid duplicating their insights here.

## 1. BEGINNINGS

The first stage in the life cycle involves the initiation of the venture. This typically requires the commitment of time and resources to found a new firm. The agents who undertake the activities leading to new venture creation are called *nascent* entrepreneurs. Launching a business successfully is a complicated process which has been the focus of a concerted research effort over the last decade. This literature has identified the use of appropriate strategies by the entrepreneur as key to making a successful start.

The two chapters in Stage 1, by Joachim Wagner (Chapter 2) and Shaker Zahra (Chapter 3) provide excellent overviews of these two topics: nascent entrepreneurs and new venture strategies. In his survey, Joachim Wagner follows contemporary empirical practice by defining a nascent entrepreneur as someone who is currently trying to start a new business; who expects to be the owner or part-owner of the new firm; and who has been active in trying to start the new firm in the past 12 months without yet having a positive monthly cash flow. Wagner’s chapter is essentially an overview of some empirical facts

about nascent entrepreneurship distilled from recent research. Wagner first summarises evidence about the incidence of nascent entrepreneurship before surveying the start-up activities of nascent entrepreneurs. He then goes on to provide a “birds-eye view” of the characteristics of nascent entrepreneurs; the factors associated with becoming a nascent entrepreneur; and the outcomes of nascent entrepreneurs, in terms of whether they start, quit or continue preparing for eventual business entry. Wagner’s chapter provides a wealth of information about nascent entrepreneurship, summarising key findings from a large, fast-growing and diverse literature which looks set to continue growing rapidly in the years ahead.

In Chapter 3, Shaker Zahra reviews the literature on new venture strategy and its implications for organizational survival, financial performance and growth. Zahra discusses competitive, cooperative and political strategies, and highlights the importance of synchronizing these strategies. A particular strength of this chapter is its ability to uncover similarities and unifying themes in apparently divergent views. In this way, the chapter succeeds in resolving some apparent contradictions by putting the different perspectives into a complementary context and identifying areas where convergence appears to be within reach. Zahra goes on to highlight the contributions of the different strategic approaches, as well as their shortcomings. He concludes by identifying several implications for future research.

The empirical and theoretical perspectives of these two chapters weave a rich and complementary tapestry of issues facing entrepreneurs early on in the venture life cycle. Particular attributes and circumstances appear to set nascent entrepreneurs apart from other people; and what happens to them after they are launched—the topic Wagner ends up discussing—is shaped (as Zahra shows) by the strategic choices that entrepreneurs make at the outset, and which they modify as their business develops. Perhaps this is the most crucial stage of the life cycle in the sense that if entrepreneurs do not make the right choices at the start, they may never survive long enough to contend with the challenges that emerge at later stages.

## 2. ASPECTS OF ENTRY AND NEW VENTURE CREATION

Information about opportunities, strategies and resources is crucial to the entry process. It appears that the entrepreneurship literature has sometimes discussed opportunities, information and venture performance in a rather passive way, asserting, for example, that some people are just innately more alert or responsive than others. Echoing the increasing policy interest in successful entrepreneurship, however, another strand of the literature is beginning to take a more proactive approach by asking whether entrepreneurship education can

purposefully augment and enhance the skills that are needed to succeed in entrepreneurship. As Lena Lee and Poh-Kam Wong remark in Chapter 4, the literature on entrepreneurship education has to date been “largely scattered and sporadic.” Lee and Wong’s chapter usefully consolidates and synthesises this literature by reviewing what we know about the extent to which entrepreneurship can be taught, and by evaluating the structure, efficacy and growth of entrepreneurship education courses. An interesting insight from this chapter is that entrepreneurship education should be tailored to the development of the venture. This recognizes that the learning needs of entrepreneurs generally change as their ventures evolve.

The findings of Herbert J. Schuetze and Heather Antecol (Chapter 5) also point to the importance of temporal changes in the propensity to venture, here with regard to immigrants. This is the first of three successive chapters that explore different economic and demographic aspects of new venture creation. Schuetze and Antecol observe that despite very different rates of self-employment in Australia, Canada and the United States, propensities to be self-employed among immigrant men in these countries tend to catch up with and then overtake those of otherwise similar indigenous citizens within 10 to 20 years after arrival. The “quality” of these immigrant enterprises, on the other hand (as measured in terms of earnings outcomes) is uneven, varying between these three countries. Schuetze and Antecol discuss the likely impact of immigration policies in these countries on the quality and quantity of immigrant start-ups—in the context of other institutional and market factors that affect venturing. Australia’s relatively rigorous “points” requirements for entry appear to have had the expected effects both in terms of promoting business start-up and facilitating favourable earnings outcomes among those who have created new ventures.

Even if entrepreneurs possess the requisite training and quality to enter the market, environmental factors have an important impact on their emergence and early performance prospects. One environmental factor that is attracting growing interest is geographical location. In Chapter 6, David Audretsch and Erik Lehmann explain how and why geography matters for new venture creation and the performance of new firms. A key element in the new geography of entrepreneurship is the existence of knowledge spillovers, which new ventures are often uniquely well placed to exploit. Audretsch and Lehmann review the extensive literature on spillovers, location and start-ups, and then go on to present some novel results derived from a new database comprising 281 publicly listed new ventures in German high-technology and knowledge-intensive industries. The chapter contains a lively blend of theoretical and empirical insights that illuminate this topical and interesting research area.

Geography is one of the factors affecting entry and exit decisions covered in Martin Carree’s review. Carree first connects the entry decision to

the multi-disciplinary entrepreneurship literature, and then reviews the last four decades of empirical research on entry and exit of firms in terms of incentives and barriers to entry and exit. To complement earlier chapters that focused on individual- or micro-level approaches to new venture creation, Carree's chapter takes a more aggregate look at these processes, including variations in entry and exit rates over time, across space and between industries.

Together, these four chapters provide comprehensive overviews of the individual, spatial and industry-specific factors that bear on the entry decision. Each of them is stamped with a strong policy emphasis. This emphasis is projected at the individual level (Lee-Wong and Schuetze), at the regional level (Audretsch and Lehmann) and at the industrial and national level (Carree). Although these policy implications are not always discussed in detail, the reader should usually be able to surmise the relevance of the material covered here to the policy community.

### 3. FINANCING VENTURES

Many ventures require finance to help them launch and grow. Finance can enable ventures to attain sufficient size to overcome the scale entry barriers mentioned in Carree's chapter and to compete effectively with incumbent firms. And, of course, external funds are often needed to finance growth later in the venture life cycle. Finance issues are the focus of four chapters in this section of the volume.

Two of the most important forms of external venture finance in developed countries are debt finance and equity finance. Robert Cressy (Chapter 8) and Christian Keuschnigg and Søren Bo Nielsen (Chapter 9) provide overviews of important aspects of formal debt and equity finance. These chapters are complemented by Colin Mason's review of informal sources of venture finance (Chapter 10) and João Pedro Azevedo's discussion of micro-finance schemes (Chapter 11) which are more commonly observed in developing countries.

Each of these chapters extends their purview beyond the traditionally narrow confines of these topics. In his comprehensive chapter, Cressy explains the structure of debt finance in practice; explores the role of collateral and the possibility of credit constraints; and assesses market and government solutions to the problem of limited credit. Considerable attention is devoted to the proposition that a correlation between entry into entrepreneurship and personal wealth signals credit constraints. In a similar fashion, Keuschnigg and Nielsen go beyond a discussion of venture capital investment to consider in depth the interface between public policy, venture capital and entrepreneurship. For his part, Mason supplements his overview of the scale and scope of the informal venture capital market with a discussion of the investment process observed in informal

venture capital, and an explanation of the new organizational formats that are emerging for angel investing. And Azevedo moves beyond institutional detail to consider the pressing issue of poverty reduction in developing countries.

The role of public policy evidently plays a prominent role in each of these chapters. In Chapter 8, Cressy discusses the possibility that start-ups face credit constraints that either prevent individuals from launching new ventures or that leave them under-capitalised if they do. Under-capitalisation can leave firms vulnerable to failure later in the life cycle if, for example, they are poorly equipped to fend off competition that intensifies as the venture seeks to grow (see also Chapter 17 of this volume). Cressy's chapter contains a detailed analysis and critique of credit constraint models before considering what markets and governments can do to mitigate them. Cressy concludes that governments are invariably better placed to resolve these problems than markets are.

In Chapter 9, Keuschnigg and Nielsen analyse the effects of public policy for venture capital. This chapter discusses the consequences of various taxes and subsidies for the rate of business creation and the quantity and quality of Venture Capitalist (VC) financed entrepreneurship. These include subsidies to start-up investment, capital gains taxation, income taxation and corporate income taxation. An important insight of this chapter is that these taxes all become relevant at different stages of the firm's life cycle. For example, the taxation of mature firms might be as important for start-ups as the direct taxation of infant companies because, by reducing the value of mature firms, the corporate tax diminishes the gains from setting up a new venture in the first place. This motivates Keuschnigg and Nielsen's use of a fairly technical multi-period model. The implications of their model are striking, suggesting that there is a quality-quantity trade-off in the promotion of new ventures, and that policies should aim more for quality than for quantity. A better quality of start-up goes hand in hand with superior incentives for entrepreneurial effort and VC support and results in lower failure rates among start-up firms.

Informal venture capital typically plays a much greater role in funding start-ups and the initial growth stages of entrepreneurial ventures than formal venture capital does. Despite this, the literature and the number of active researchers in this field remain small. In Chapter 10, Colin Mason offers a definitive overview of the attitudes, behaviors and characteristics of business angels, their economic performance, and government efforts to expand the supply of informal venture capital. The following government interventions receive particular attention: business angel networks and tax incentives. However, the efficiency of these interventions is not yet well understood. Future researchers might fruitfully attempt to apply modelling techniques such as those used in Chapter 9 to the informal venture capital sector to better understand the scope for public policy in this domain.

In Chapter 11, Azevedo explores market rather than government solutions to problems of borrowing constraints in developing countries. These market solutions take the form of microcredit provided by microfinance institutions (MFIs). MFIs are an increasingly important source of funds in developing countries where access to formal credit markets is often limited, for reasons that Azevedo explains. Looking only at credit services, MFIs account for about 33% of overall loans in the countries under study. As Azevedo points out, this is an impressive accomplishment given that MFIs as institutions are relatively young and small. After describing the context of microenterprises, Azevedo explains how and why microentrepreneurs benefit from the provision of these financial services. Azevedo also addresses several other important issues in his chapter, including the funding mechanisms and financial self-sustainability of MFIs, their impact, and policy implications of using microfinance as a microenterprise promotion tool.

#### 4. VENTURE DEVELOPMENT I: PRIVATE SECTOR ISSUES

Once an entrepreneur has identified their business opportunity, formulated their strategy and obtained the necessary finance, he or she launches his or her venture. The entrepreneur has to decide what to produce and in what quantity, and what factors of production to utilize in order to achieve his or her goal. This is a crucial aspect of venture development, as unwise decisions at this stage can have long-term deleterious consequences for the development and even survival of the venture.

Two pertinent issues are explored in this part of the volume. In Chapter 12 the entrepreneur's production decision is discussed. It is argued in that chapter that a production function framework can link together several salient issues in entrepreneurship, including entrepreneurial psychology, innovation, entrepreneurship education and female entrepreneurship. A variety of aspects of production that are specific to entrepreneurs and entrepreneurship are also discussed, including entrepreneurial ability, the entrepreneur's choice of production function, knowledge spillovers and "entrepreneurship capital." The chapter closes with some suggestions for future research, including ways of consolidating the linkages that are explored here.

While some dynamic aspects of venture development are discussed in Chapter 12 (notably the entrepreneur's decision to change technology), that chapter does not address directly the question of growth. Two chapters in this volume treat this issue, the first being Chapter 13, by Per Davidsson, Leona Achtenhagen and Lucia Naldi (the second, by Marc Cowling, is connected to survival so is treated at a later stage of the volume). These authors discuss a comprehensive range of issues relating to venture development, including how



growth is measured, the factors that facilitate and hinder it, stages and transitions in the growth process and several implications of growth. They also discuss different ways that ventures can grow, including growth by acquisition and growth by expansion into international markets. The authors argue that current research on growth has largely ignored the particularities of small firms and has in turn been under-researched by small business scholars.

Clearly, production decisions affect ventures' survival and growth prospects. Hence the issues covered in Chapters 12 and 13 are indirectly linked. Most of the discussion in these chapters centers on private returns that are captured by the entrepreneurs themselves. However, entrepreneurs' decisions and venture development outcomes carry wider socio-economic implications as well. Hence the next part of this handbook addresses social issues arising from venture development.

## 5. VENTURE DEVELOPMENT II: SOCIAL ISSUES

The two previous chapters discussed the challenges facing entrepreneurs trying to maximize their private returns net of their costs. Attention turns in the present section to some broader social issues arising from venture development. These issues are varied and numerous, and we only have space to treat two, specifically social entrepreneurship and entrepreneurship among disadvantaged groups, namely women, minorities and the less educated. To some extent, these topics tie in with gender and ethnicity issues that were identified in Chapter 13.

Helen Haugh discusses nonprofit social entrepreneurship in Chapter 14. This topic currently occupies a rather peripheral position on the edge of entrepreneurship research; her chapter therefore does a major service by assembling in one place the scattered and occasionally transient literature relating to it. This is all the more important as research into social entrepreneurship is beginning to burgeon, perhaps reflecting the growing social and economic importance of this sector. As Haugh points out, by the late 1990s the nonprofit sector accounted for aggregate expenditure of \$1.3 trillion, representing around 5% of Gross Domestic Product in many countries and employing over 4% of the economically active population. Haugh uses a supply and demand framework to understand the nature of social entrepreneurship, and she examines the characteristics and motivations of social entrepreneurs and the process of social entrepreneurship, including social venture creation and performance measurement. She emphasizes the increasing pressure on nonprofit organisations to adopt more enterprising strategies. An impressively wide array of material covering a variety of topics, issues and perspectives is combined in a truly interdisciplinary, information-packed chapter.

Another social issue that engages policy makers is entrepreneurship among disadvantaged groups, including women, ethnic minorities and less-educated workers. The creation and development of sustainable businesses, it is hoped, can help members of these disadvantaged groups escape discrimination and poverty. Robert Fairlie explores this issue in Chapter 15, where he first documents and then seeks to explain variations in business ownership rates by gender, ethnicity and education in Britain, the USA, Canada and other countries of the OECD. Fairlie splits self-employment rates into entry and exit components and so forms a bridge between the first and last stages of the venture life cycle. As he observes, venture development is conditional on both entry and continuation (survival) so it is helpful to analyse entry and exit separately. This distinction also turns out to be useful for understanding why business ownership rates are so low for members of some of these disadvantaged groups. Blacks in particular have both lower entry rates and higher exit rates than whites. The determinants of entry and exit are found to be different, although education (Chapter 17) and access to bank finance (Chapter 8) tend to affect both. That much of the underlying variation in these diverse entry and exit patterns remains unexplained only highlights the urgent need for further research in this area.

As should be clear from the preceding chapters, venture creation and development carry economic and social implications which extend beyond private benefits that can be appropriated by entrepreneurs. Successful venture development may therefore advance the interests of society as a whole, not just a small enclave of business owners.

## 6. VENTURE PERFORMANCE AND HARVESTING

The final stage of the venture development process brings together several phenomena, including survival, profitability and harvesting. The remaining chapters of the volume deal with each of these topics.

Pressures to survive and grow are often intense for small entrepreneurial firms. As Marc Cowling points out in Chapter 16, it is helpful to take account of survival when analyzing venture performance—not least in order to avoid problems of survival bias entailed by studying only surviving ventures. In his chapter, Cowling accordingly explores the two issues together. He draws on the available evidence to identify the determinants of survival and growth. His overview emphasizes a range of factors, including human capital, personal characteristics, business characteristics and macroeconomic conditions. Cowling also spends some time discussing the impact of public unemployment assistance programs on survival rates, and the effects of entrepreneurs' competencies and strategic decisions on venture growth. These findings are likely to be of

particular interest to policy makers and entrepreneurs. Cowling's treatment of growth complements that given by Davidsson and co-authors in Chapter 13. His chapter is positioned at a more aggregate level and provides more of a policy discussion than Davidsson et al's; it also treats different material and is "narrower but deeper" in its treatment than Davidsson et al. Cowling also provides helpful summary tables comparing previous studies which facilitates an in-depth treatment of the determinants of these phenomena.

Of course, survival and growth are two widely used measures of venture performance; but other measures exist too, the most common being profitability. In Chapter 17, Mirjam van Praag reviews the impact of human and financial capital on profitability with an emphasis on the role of formal education. She urges entrepreneurship scholars to use state-of-the art econometric techniques to measure the returns from schooling in entrepreneurship. She goes on to summarize some of the recent empirical findings from studies which point to higher returns for entrepreneurs than for employees. Van Praag also extends Robert Cressy's review of the empirical credit constraint entrepreneurship literature by asking how credit constraints affect profitability and survival (rather than participation in entrepreneurship, which was the focus of Cressy's chapter in this regard). She claims that recent research shows that financial constraints at start-up hinder entrepreneurial performance later on in the life cycle. These findings carry obvious policy implications which are briefly discussed at the end of her chapter.

The life cycle of entrepreneurial ventures inevitably terminates with closure. Closure can be involuntary when it is often referred to as "failure"; or it can be voluntary, when it is known as "harvesting." Involuntary closure is discussed in several places in the volume, chiefly in Chapters 8, 15, 16 and 17. To complete the picture, the final chapter of this volume, Chapter 18 by Maks Tajnikar, Petra Došenovič Bonča and Lidija Zajec, treat harvesting. As these authors point out, although harvesting is often associated with an orderly withdrawal from a venture, the process is not always related to exit. They consider five forms of harvesting: free cash flow and dividend payouts; public offerings (IPOs); trade sales, buybacks, management and employee buyouts and buy-ins; acquisitions, mergers and takeovers; and liquidations, bankruptcies, sale of assets and write-offs. The last of these forms is closely related to involuntary exit and so the term "harvesting" as used by these authors is very broad. The chapter classifies into four groups the factors that determine the entrepreneur's and investor's choice of these harvesting forms. In the process, Tajnikar and co-authors span a wide range of literature that admirably draws the volume, as well as the venture's life cycle, to a close.

## 7. CONCLUSION

This volume has two goals. These are to assemble several chapters on important issues in entrepreneurship which are of independent interest, and to do so within an over-arching framework of the venture life cycle that is illuminating and thought provoking. A framework as expansive as this necessarily draws on a range of different disciplines and cross-national perspectives. Of course, plenty more research needs to be done at each stage of the venture life cycle, and the chapters that comprise this collection have invariably tried to prioritize particular questions on which future research effort should concentrate. It is with a renewed plea for further research grounded in the existing literature that the reader is welcomed to this volume.

## Stage 1: Beginnings

JOACHIM WAGNER

*Institute of Economics, University of Lueneburg, Campus 4.210, D-21332 Lueneburg, Germany*  
*E-mail: Wagner@uni-lueneburg.de*

## 2. Nascent Entrepreneurs

### 1. WHAT IS A NASCENT ENTREPRENEUR?<sup>1</sup>

The creation of a new venture is a process. Following Reynolds and White (1997, p. 6) and Reynolds (2000, p. 158ff.), this process, analogous to biological creation, can be considered to have four stages (conception, gestation, infancy and adolescence), with three transitions. The first transition begins when one or more persons start to commit time and resources to founding a new firm. If they do so on their own and if the new venture can be considered an independent start-up, they are called *nascent entrepreneurs*. If the entrepreneurship occurs within the context of an existing organization, they are considered to be nascent intrapreneurs. The second transition occurs when the gestation process is complete and when the new venture either starts as an operating business, or when the nascent entrepreneurs abandon their effort and a stillborn happens. The third transition is the passage from infancy to adolescence—the fledgling new firm’s successful shift to an established new firm.

This chapter deals with the first two stages and the first two transitions of this process and with their main actors—nascent entrepreneurs. This means that we will neither look at nascent intrapreneurs, nor will we deal with the survival (or not) and growth pattern of active new firms. And we will not look at those who just state that they would prefer being self-employed over being an employee—a group which can be labeled latent entrepreneurs (Blanchflower, 2004; Blanchflower et al., 2001). Instead, we will focus on people who are currently taking explicit steps to start a new business. To fix ideas and following the definition used in the Panel Study of Entrepreneurial Dynamics (PSED) (Reynolds, 2000, p. 170f.; Shaver et al., 2001; Gartner and Carter, 2003, p. 203f.; Gartner et al., 2004; Reynolds et al., 2004a) and in the

Global Entrepreneurship Monitor (GEM) (Reynolds et al., 1999, 2000, 2001, 2002a, 2004b; Acs et al., 2005), a nascent entrepreneur is defined as a person who is now trying to start a new business, who expects to be the owner or part owner of the new firm, who has been active in trying to start the new firm in the past 12 months and whose start-up did not yet have a positive monthly cash flow that covers expenses and the owner-manager salaries for more than three months.

One advantage of this definition is that it makes clear in an operational way who is a nascent entrepreneur. From the answers to a number of survey questions we can decide whether a person is to be considered a nascent entrepreneur or not. If, for example, someone argues that he recently decided to become self-employed in the future, but did not take active steps of the kind outlined above, he is not. The decision to become self-employed comes first—taking first steps comes next (sometimes). The status of being a *nascent* entrepreneur or not according to the definition used here can be identified empirically. Given the various definitions of who is an *entrepreneur* (see Davidsson, 2004, ch. 1) it is evident that it depends on the specific definition chosen whether a person who is considered a nascent entrepreneur is viewed as an entrepreneur, too.

Using the definition of a nascent entrepreneur outlined above, the rest of the chapter will use the existing economics literature to discuss the following questions: How many nascent entrepreneurs are there (Section 2)? What do nascent entrepreneurs do (Section 3)? Who are the nascent entrepreneurs (Section 4)? What makes a nascent entrepreneur (Section 5)? What happens to nascent entrepreneurs and why (Section 6)? Section 7 concludes.

## 2. HOW MANY NASCENT ENTREPRENEURS ARE THERE?

Given that newly founded firms are important for the economic development of nations and regions (see Carree and Thurik, 2005, for a short overview of the recent literature on the role of entrepreneurship for economic growth and development) and that nascent entrepreneurs are by definition important for the foundation of new firms, information about nascent entrepreneurs is important for understanding crucial aspects of the economy. This information, however, can not be found in publications from official statistics. Some ten years ago, therefore, we knew next to nothing about nascent entrepreneurs. The situation improved considerably when results from two pioneering studies—the Wisconsin Entrepreneurial Climate Study conducted in spring 1993 and a national pilot study for the U.S. done in October/November 1993—were published (see Reynolds and White, 1997). Furthermore, in the U.S., the Panel Study of Entrepreneurial Dynamics (PSED) that started in 1998 now is a representative national database on the process of business formation (Reynolds, 2000; Reynolds et al., 2002b, 2004a; Gartner et al., 2004).

In 1999 the Global Entrepreneurship Monitor (GEM) project was started (Reynolds et al., 1999, 2000, 2001, 2002a, 2004b; Acs et al., 2005). At the heart of this international project are representative surveys of the adult population in the participating countries that use an identical questionnaire to measure various aspects of entrepreneurial activity. The share of nascent entrepreneurs in the population is measured by asking the interviewees a set of questions that closely follows the definition of a nascent entrepreneur given in Section 1. Thirty one countries participated in the 2003 wave of GEM and some 100,000 adults were interviewed. Table 2-1 reports the share of nascent entrepreneurs computed from these surveys (together with the lower and upper bounds of the 95% confidence interval for the point estimates).

From Table 2-1 it is evident that there are millions of nascent entrepreneurs. Using the figures reported for the share of nascent entrepreneurs in Table 2-1 and the numbers for the total population 18–64 years old from the GEM 2003 executive report (Reynolds et al., 2004b, p. 16), one calculates that in 2003 there were some 14.689 million nascent entrepreneurs in the U.S., 1.843 million in Germany and 1.271 million in the U.K.

Table 2-1 reveals one more striking fact: The share of nascent entrepreneurs differs widely between countries. While in Venezuela in 2003 one in five adults was a nascent entrepreneur, we found one in twelve in the U.S., one in 29 in Germany and the U.K. and one in 111 in France. Given that the shares are point estimates based on (representative) samples, the differences between the numbers reported in column two of Table 2-1 for two countries are not always statistically significantly different from zero at a usual error level (as can be seen from the lower and upper bounds of the 95% confidence intervals in columns 3 and 4)—consider, for example, the reported shares for New Zealand and the U.S., or Finland and Ireland. However, it is evident that there are many differences which are both statistically significant and large in an economic sense—just compare the U.S. with the U.K. and Germany and Germany with its neighbor countries France and the Netherlands. It should be noted in passing that similar differences in the share of nascent entrepreneurs have been found between regions in Germany (Wagner and Sternberg, 2004).

How can these differences in the share of nascent entrepreneurs across space be explained? What makes a country more or less entrepreneurial? Using data for 36 countries participating in the Global Entrepreneurship Monitor 2002, van Stel et al. (2003) investigate this question employing four empirical approaches: First, they hypothesize nascent entrepreneurship to be a function of the level of economic development of a country, using per capita income as an indicator. Second, they test for the influence of Porter's "Innovative Capacity Index" built from information on, among other things, the proportion of scientists and engineers in the workforce, intellectual property protection and R&D tax credits for the private sector, the presence of suppliers of specialized



TABLE 2-1 *Share of nascent entrepreneurs in the adult population (18–64 years) in 2003*

| Country      | Share of nascent entrepreneurs | Lower bound of 95% confidence interval | Upper bound of 95% confidence interval |
|--------------|--------------------------------|--|--|
| Venezuela    | 0.192                          | 0.174                                  | 0.210                                  |
| Uganda       | 0.148                          | 0.125                                  | 0.170                                  |
| Argentina    | 0.124                          | 0.109                                  | 0.140                                  |
| Chile        | 0.109                          | 0.095                                  | 0.124                                  |
| New Zealand  | 0.093                          | 0.079                                  | 0.107                                  |
| U.S.         | 0.081                          | 0.075                                  | 0.087                                  |
| Iceland      | 0.073                          | 0.061                                  | 0.085                                  |
| Australia    | 0.066                          | 0.055                                  | 0.077                                  |
| Brazil       | 0.065                          | 0.054                                  | 0.076                                  |
| Ireland      | 0.051                          | 0.040                                  | 0.062                                  |
| Canada       | 0.051                          | 0.040                                  | 0.061                                  |
| Spain        | 0.044                          | 0.039                                  | 0.049                                  |
| Switzerland  | 0.043                          | 0.034                                  | 0.053                                  |
| China        | 0.043                          | 0.033                                  | 0.053                                  |
| Finland      | 0.041                          | 0.030                                  | 0.051                                  |
| Norway       | 0.040                          | 0.030                                  | 0.050                                  |
| Germany      | 0.035                          | 0.030                                  | 0.040                                  |
| U.K.         | 0.034                          | 0.031                                  | 0.037                                  |
| Denmark      | 0.031                          | 0.023                                  | 0.038                                  |
| Singapore    | 0.030                          | 0.022                                  | 0.038                                  |
| Slovenia     | 0.030                          | 0.022                                  | 0.038                                  |
| Greece       | 0.029                          | 0.022                                  | 0.037                                  |
| Belgium      | 0.028                          | 0.021                                  | 0.035                                  |
| South Africa | 0.027                          | 0.021                                  | 0.034                                  |
| Italy        | 0.020                          | 0.014                                  | 0.027                                  |
| Sweden       | 0.020                          | 0.014                                  | 0.027                                  |
| Croatia      | 0.018                          | 0.011                                  | 0.024                                  |
| Netherlands  | 0.017                          | 0.012                                  | 0.022                                  |
| Hong Kong    | 0.017                          | 0.011                                  | 0.023                                  |
| Japan        | 0.014                          | 0.008                                  | 0.019                                  |
| France       | 0.009                          | 0.004                                  | 0.013                                  |

Source: Global Entrepreneurship Monitor 2003 (data provided by Rolf Sternberg).

research and training, the quality of scientific research institutions and the availability of venture capital (for details, see Porter and Stern, 2002). Third, they take an eclectic stand and link nascent entrepreneurship to a portfolio of determinants including economic and noneconomic conditions, such as technology, demography, culture and institutions. Fourth, they combine the approaches mentioned before in a single empirical model.

Both for the relationship of the share of nascent entrepreneurs with per capita income and with innovative capacity, van Stel et al. (2003) find a u-shaped relationship. Rising levels of economic development and innovative

capacity go along with a declining share of nascent entrepreneurs in the adult population up to a certain level and then start to rise again as per capita income or the index of innovative capacity increases still further. Using the empirical model based on the eclectic approach, they start with a set of twelve exogenous variables and apply a stepwise procedure to end with four determinants (the sign of the estimated regression coefficient is given in brackets): A variable measuring the stock of incumbent business owners (+), the innovative capacity index (–) social security costs as percent of GDP (–) and a dummy variable indicating whether a country has been a communist country in the past or not (–). In the full model, combining the other three approaches, the positive relationship with the stock of incumbent business owners, the negative impact of being a former communist country and the u-shaped relationship with the innovative capacity index still hold (while the u-shaped relationship with per capita income is no longer statistically significant at a conventional level); the regression coefficient of the social security costs variable remains negative, but is no longer statistically significant at a conventional level.

The authors themselves point out three limitations of their study: It is based on cross-section data for one moment in time only; it does not disaggregate by sector of activity (industry vs. services, etc.) nor does it make a distinction between “necessity entrepreneurship” (which refers to a situation where people are involved in entrepreneurship activities because they have no better choices for work) and “opportunity entrepreneurship” (where people are pursuing a business opportunity or personal interest while other choices to earn a living are open to them); and it assumes that the same empirical model is appropriate for countries as different as the U.S., Russia and Brazil. Furthermore, the stepwise approach used might be expected to end up in an empirical model that is tailor-made for the data set at hand.

In an empirical investigation that has a focus on the role of post-materialism as a cultural factor influencing cross-country differences in total entrepreneurial activity (defined as the share of nascent entrepreneurs plus the share of people who are owner-managers of a business less than 42 months old) Uhlaner and Thurik (2004) report estimates from an empirical model regressing the share of nascent entrepreneurs on five variables (see their Table 2, column 7). The study is based on data from 28 countries which is a subset of the countries that participated in the Global Entrepreneurship Monitor 2002 and, therefore, a subset used in the study by van Stel discussed above. They find a weakly significant positive relationship with an index of life satisfaction and a highly significant relationship with the gross enrollment ratio in secondary education; the estimated coefficients of the variables measuring post-materialism, per capita income and the gross enrollment ratio in tertiary education are statistically insignificant at any conventional level. Given that the investigation of cross-country differences in the share of nascent entrepreneurs

is not at the center of the study and the limitations of the study (which are similar to those mentioned in the context of the van Stel et al. study), the results should not be expected to shed much light on the topic considered here.

The two pioneering studies by van Stel et al. (2003) and by Uhlaner and Thurik (2004) are (to the best of my knowledge) the only large-scale empirical investigations looking at cross-country differences in the share of nascent entrepreneurs (see also Welter, 2001 for a comparison of Germany, the Netherlands, Sweden and the USA; for further studies on cross-country differences in total entrepreneurial activities—defined as the share of nascent entrepreneurs plus the share of people who are owner-managers of a business less than 42 months old—see Verheul and Thurik, 2003 and Verheul and Thurik, 2003). A limitation of these studies is the likely endogeneity of some of the covariates. Furthermore, it is an open question and one well worth future research efforts, whether the findings in these studies can be replicated for different samples of countries and for different periods and what is the role played by other factors not investigated hitherto.

### 3. WHAT DO NASCENT ENTREPRENEURS DO?

What activities are nascent entrepreneurs involved in when they are actively engaged in creating a new venture of their own? The only way to find out is to ask them and this has been done in the U.S. in the Wisconsin Entrepreneurial Climate Study conducted in Spring 1993, in a national pilot study for the U.S. done in October/November 1993 (Reynolds, 1997; Reynolds and White, 1997) and in the Panel Study of Entrepreneurial Dynamics (PSED, formerly Entrepreneurial Research Consortium/ERC) that started in 1998 (Reynolds, 2000; Reynolds et al., 2001, 2002a; Gartner and Carter, 2003). Furthermore, we have evidence from ERC-based surveys conducted in Norway (Alsos and Ljunggren, 1998) and in Canada (Diochon et al., 2001).

In the order of “popularity” among the respondents in the U.S. sample of 1993, the following start-up activities were reported by at least one third of the nascent entrepreneurs (Reynolds, 1997, p. 452; Reynolds and White, 1997, p. 41): Serious thought about business; looked for facilities/equipment; initiated savings to invest; invested own money in the new firm; organized start-up team; written business plan; bought facilities/equipment; sought financial support; license, patent, permits applied for; developed first model or prototype; received money from sales. About 95% of the nascent entrepreneurs indicated two or more start-up behaviors; the median number of steps taken was seven. Using a similar (but not identical) list of activities, Diochon et al. (2001) report similar results from interviews with some 120 nascent Canadian entrepreneurs performed in 2000: respondents are engaging in multiple activities and the

most intensely pursued are: defining market opportunities; personally investing money in the venture; purchasing raw materials, inventory, supplies or components; generating sales revenue; and marketing, promotional efforts. Looking at gender differences in start-up activities among 114 male and 35 female Norwegian nascent entrepreneurs interviewed in 1997, Alsos and Ljunggren (1998) find few differences between male and female nascents—among others, a smaller proportion of the women than of the men reported having prepared a business plan and hired employees.

Evidence on the “first behavior” of nascent entrepreneurs based on the interviews from the PSED is reported by Gartner and Carter (2003, p. 203f.). According to their findings, 57% of the 715 nascent entrepreneurs “spent a lot of time thinking about starting business” first, followed by 16% who “took classes or workshops on starting business,” 15% “saving money to invest in business,” 14% “invested own money in business” and 12% “developed model or procedures for product/service.” The authors list 21 more start-up behaviors that occurred first among less than 10% of the nascent entrepreneurs. Carter and Kolvereid (1998) compare first activities between male and female nascent entrepreneurs in the U.S. and in Norway and they find variation across both gender and country.

Unfortunately, we do not have comprehensive and comparable evidence on the set of activities nascent entrepreneurs are involved in and on the timing of these events, for a large number of countries, because this is a topic that has not yet been investigated in the Global Entrepreneurship Monitor project. From the evidence we have on start-up activities it is clear that there is neither a fixed set of events (although some events are more common than others) nor a uniform sequence. The industry, the region and personal factors (like gender, skills and financial reserves of the nascent entrepreneurs) all matter in determining what a nascent entrepreneurs does and when.

#### 4. WHO ARE THE NASCENT ENTREPRENEURS?

Are nascent entrepreneurs different from the rest of the adult population and is there a typical nascent entrepreneur with a typical set of characteristics? Table 2-2 reports the relationship between the prevalence rate of nascent entrepreneurs and selected personal characteristics and attitudes. This evidence is based on the (weighted) data from the 29 countries that took part in the Global Entrepreneurship Monitor in 2001 (Reynolds et al., 2001, p. 32).

According to Table 2-2, the share of nascent entrepreneurs in the total population covered by the surveys is much higher for men than for women and it declines with age; it is more than twice as high for those who know an entrepreneur than for those who do not and more than three times higher for

TABLE 2-2 *Impact of selected factors on nascent entrepreneurship (29 GEM countries, 2001)*

| Factor                                       | Share of nascent entrepreneurs in selected groups of people |
|--|---|
| <i>Gender:</i>                               |   |
| Men  | 9.3%  |
| Women  | 4.2%  |
| <i>Age:</i>                                  |   |
| 18–24 years old                              | 8.0%  |
| 25–34 years old                              | 7.9%  |
| 35–44 years old                              | 7.5%  |
| 45–54 years old                              | 5.2%  |
| 55–64 years old                              | 4.5%  |
| <i>Contact with entrepreneurs:</i>           |   |
| Know an entrepreneur: Yes                    | 11.6%   |
| Know an entrepreneur: No                     | 5.1%  |
| <i>Perception of business opportunities:</i> |   |
| Good opportunity for business: Yes           | 14.5%   |
| Good opportunity for business: No            | 4.3%  |
| <i>Business skills:</i>                      |   |
| Have skills to start a business: Yes         | 13.8%   |
| Have skills to start a business: No          | 2.4%  |
| <i>Fear of failure:</i>                      |   |
| Failure fear NOT a problem: Yes              | 8.5%  |
| Failure fear NOT a problem: No               | 4.1%  |
| <i>Family's economic future:</i>             |   |
| Family future looks: Better                  | 10.5%   |
| Family future looks: Same                    | 4.4%  |
| Family future looks: Worse                   | 3.3%  |
| <i>Country's economic future:</i>            |   |
| Country future looks: Better                 | 8.6%  |
| Country future looks: Same                   | 5.1%  |
| Country future looks: Worse                  | 6.3%  |
| <i>Educational attainment:</i>               |   |
| Graduate program experience                  | 5.4%  |
| Beyond secondary school                      | 7.6%  |
| Secondary school degree                      | 8.4%  |
| Not completed secondary school               | 5.7%  |
| <i>Labor force status:</i>                   |   |
| Working full or part time                    | 8.4%  |
| Not working: Homecare, unemployed            | 4.3%  |
| Not in labor force: Retired, student         | 3.4%  |
| <i>Relative household income:</i>            |   |
| HH income in upper third for country         | 7.9%  |
| HH income in middle third for country        | 6.9%  |
| HH income in lower third for country         | 6.1%  |

Source: Global Entrepreneurship Monitor 2001 Summary Report (Reynolds et al., 2001, p. 32).

those who perceive a good opportunity for business compared to those who do not; the presence of business skills increases the share by a factor of nearly 6. Fear of failure matters—the share of nascent entrepreneurs is twice as high among those who fear failure as those who do. The better the family future looks, the higher is the prevalence rate of nascent entrepreneurs; the link with the perception of the country's economic future, however, is nonmonotonic with the lowest share of nascents among those who state that the country's future looks the same as today. As regards educational attainment, the share of nascents is lowest for those at the top and at the bottom end and considerably higher in between. Nascent entrepreneurs are more often found among individuals who are working full or part time than among those who are not working or are not in the labor force. The higher the household income, the higher is the prevalence rate of nascent entrepreneurs. This evidence from the Global Entrepreneurship Monitor project shows that certain types of individuals are more likely to be involved in creating a new venture, but that individuals from all categories are involved to some extent.

Although the evidence reported in Table 2-2 reveals important facts about nascent entrepreneurs two shortcomings are evident.

First, a look at the (weighted) average of data from 29 countries in one year is a bird's eye view—a closer look at data for single countries (or regions inside countries) and several years will demonstrate important differences across both space and time. Fortunately, there are detailed annual country reports for each country which took part in the Global Entrepreneurship Monitor project and most of these reports are available free of charge from the project's homepage ([www.gemconsortium.org](http://www.gemconsortium.org)). Furthermore, comprehensive descriptive information on nascent entrepreneurs in selected countries are available from other sources, too—for the U.S. (see evidence based on the Panel Study of Entrepreneurial Dynamics reported in Reynolds et al., 2002b, 2004a), Canada (Diochon et al., 2001), Sweden (Delmar and Davidsson, 2000), for Germany as a whole (Welter, 2001) and for selected regions in Germany (Bergmann et al., 2002; Lückgen and Oberschachtsiek, 2004). This provides researchers interested in a specific country, or in inter-country comparisons, with a rich set of information; and it offers the possibility to augment the bird's eye view given in Table 2-2 by views through a looking glass.

Second, the empirical evidence reported in Table 2-2 is only descriptive in nature and it does not reveal the extent to which the various factors considered are interrelated. To give just one example, consider the relationship between gender and nascent entrepreneurship on the one hand and between labor force status and nascent entrepreneurship on the other hand. Men are more than twice as often involved in creating new ventures than women and so are people who are working full or part time compared to those who are not working or are not in the labor force. Given that the share of men who are in paid employment

is much higher than the share of women, what is the *ceteris paribus* effect of being male and of working full or part time, on the propensity of being a nascent entrepreneur? Descriptive bivariate comparisons cannot reveal this. Multivariate analyses that tackle this topic are reviewed in the next section.

## 5. WHAT MAKES A NASCENT ENTREPRENEUR?

Empirical investigations of the *ceteris paribus* impact of individual (and other) characteristics and attitudes on the propensity to become a nascent entrepreneur are usually—either explicitly or implicitly—based on a theoretical framework that can be outlined as follows.

Consider a utility-maximizing individual that has the choice between paid employment and self-employment (taking the decision to participate in the labor market as given). This person will choose the self-employment option if the discounted expected life-time utility from self-employment ( $DELU_s$ ) is higher than that from paid employment ( $DELU_p$ ). The difference  $N_i$  between  $DELU_{si}$  and  $DELU_{pi}$ ,

$$N_i = DELU_{si} - DELU_{pi} \quad (1)$$

is therefore crucial for the decision of individual  $i$  and it will choose self-employment if  $N_i$  is positive.  $DELU_{si}$  and  $DELU_{pi}$  are determined by the expected monetary and nonmonetary returns from self-employment and paid employment according to the utility function of the person and the individual's discount rate. Higher returns lead to higher values of  $DELU$ .

The expected monetary and nonmonetary returns from both types of employment depend on variables that are either endowments of the individual  $i$  (like age, a university degree or the degree of risk-aversion) or other relevant variables (like characteristics of the region a person lives in). All these variables are summarized in a vector  $x_i$ . Given that  $N_i$  depends on  $DELU_{si}$  and  $DELU_{pi}$  and  $DELU_{si}$  and  $DELU_{pi}$  depend on the monetary and nonmonetary returns,  $N_i$  can be written as a function of  $x_i$ :

$$N_i = N_i(x_i). \quad (2)$$

Elements of  $x_i$  that have a more positive or less negative impact on  $DELU_{si}$  than on  $DELU_{pi}$  increase  $N_i$  (and vice versa). Given that the expected monetary and nonmonetary returns from both types of employment, the utility function and the discount rate of an individual are unknown to an observer, we cannot observe  $N_i$ . Therefore, we cannot test directly whether an individual characteristic or attitude (say, a university degree or a high degree of risk aversion) has a positive impact on  $N_i$  or not. If, however,  $N_i$  is greater than the

critical value zero, according to our theoretical framework, a person will choose to become an entrepreneur and the decision to do so or not is observable.

Empirical models that investigate the *ceteris paribus* influence of the elements of  $x_i$  on the probability that a person is a nascent entrepreneur use this known decision. In these models, the dummy variable indicating whether a person is a nascent entrepreneur or not is regressed on a set of exogenous variables comprising characteristics and attitudes of the individual and on other variables considered as relevant for this decision. Given the dichotomous nature of the endogenous variable these empirical models are estimated by (variants of) logit or probit and the empirical approach can be labeled a reduced form logit (or probit) approach.

Note that by focusing on the factors affecting the decision to become self-employed, as opposed to remaining in paid employment, instead of looking at differences in the probability that people are self-employed rather than employees, one avoids confounding entry and survival effects: The probability of being self-employed at a point in time depends on the probability of switching into self-employment in the past and then surviving as a self-employed until the time of the survey (see Parker, 2004, p. 25f).

While there is a large empirical literature on the *ceteris paribus* impact of personal and other variables on the probability of being an “adult” entrepreneur versus a paid employee (surveyed in Parker, 2004, Chap. 3), econometric investigations that ask what makes a nascent entrepreneur are scarce. One group of these studies deals with the more general question what makes a “typical” nascent entrepreneur, attempting to identify factors that are statistically significant for the decision to create a new venture or not. A number of econometric investigations tackle more specific issues (like gender differences in the propensity to become a nascent entrepreneur, or the role of young and small firms as hothouses for nascent entrepreneurs). These two groups of studies are reviewed in turn.

In a pioneering study, Reynolds and White (1997, p. 52ff.) and Reynolds (1997) use the data from a national pilot study for the U.S. done in October/November 1993 (mentioned above) to estimate in a first step logistic regression models predicting nascent entrepreneurs. The forward stepwise and backward stepwise procedures applied lead to slightly different “optimal models,” but three characteristics are statistically significant in both cases: age (with a negative impact) and self-employment and divorce, both of which increase the tendency. A number of other factors are present in one or the other variants of the empirical model. To consider the potential impact of interaction among the various factors, in a second step a variant of the Automatic Interaction Detection (AID) technique is applied. This leads to the identification of subgroups in the adult population where many, few or no nascent entrepreneurs can be found. For instance, 69% of new firm start-ups



are provided by 17% of the adult population: people aged 25 to 34 that are self-employed, unemployed or students and those with employment and more than a high school degree.

Further evidence for the U.S. for the determinants of the decision to become a nascent entrepreneur is reported by Kim et al. (2003) based on data from the Panel Study of Entrepreneurial Dynamics (PSED). From the results of logistic regressions they conclude that (contradicting the expectations of liquidity constraint theory) financial resources are not significantly associated with becoming a nascent entrepreneur, while several human capital variables (some college or college graduate, full-time work experience, previous start-up experience, current self-employment and the percentage of relatives who are entrepreneurs), age, being male and black or Hispanic (compared to white) all have a significant positive impact.

Comparable results for other countries are scarce. Delmar and Davidsson (2000) use an approach quite similar to the one adopted by Reynolds and White (1997) and Reynolds (1997) to look at Swedish data. Among other factors, they find a negative impact of age and positive effects of being male, having self-employed parents, education, being self-employed and having experience in management on the probability of becoming a nascent entrepreneur (see also the results from logistic regression reported in Davidsson and Honig (2003), Table 1). Using data from the first wave of the Global Entrepreneurship Monitor (GEM) for Germany collected in 1999, Sternberg (2000, p. 58f.) estimates a logit regression to investigate the *ceteris paribus* impact of age, gender, living in western or eastern Germany, size of city, education, household income and number of persons living in the household on the probability of becoming a nascent entrepreneur. He finds a strong positive effect of being male and a negative effect of being more than 54 years old.

What do we learn from these studies that attempt to identify factors that are important for becoming a nascent entrepreneur or not? In my view, not a great deal. The most important reason for this pessimistic view is that we do not have evidence from numerous studies covering many different countries and applying identical (or at least highly similar) empirical models to different data sets. Therefore, a promising strategy for further research might be the coordination of an international research project that brings together experts from many countries who agree on a common empirical methodology to be applied to comparable data sets like those from the GEM project (for a role model, see the project on regional differences in new firm formation described in Reynolds et al., 1994). From such a project we can learn a lot about what makes a nascent entrepreneur and how and why determinants differ across space and time.

Besides the papers that try to answer the question what makes a “typical” nascent entrepreneur and identify factors that are statistically significant for

the decision to create a new venture or not, several econometric investigations tackle more specific issues related to nascent entrepreneurship. This literature is reviewed below, starting with papers that focus on the *ceteris paribus* impact of one specific personal characteristic and followed by studies that investigate the *ceteris paribus* impact of elements of the environment a person lives and works in.

*Gender:* In western industrialized countries, men are on average more than twice as active in entrepreneurship as women. Little is known about precisely why this is the case. Using data from the Regional Entrepreneurship Monitor (REM) Germany, a recent representative survey of the adult German population described in detail in Lückgen and Oberschachtsiek (2004), Wagner (2004a) estimates an empirical model for the decision to become self-employed to test for differences between women and men in the *ceteris paribus* impact of several characteristics and attitudes, taking the rare events nature of becoming an entrepreneur into account. Furthermore, a nonparametric approach using Mahalanobis distance matching of men and women who are as similar as possible is used to investigate the difference in the propensity to become self-employed by gender. The core finding of this empirical exercise is that fear of failure as a reason not to start a business has a much smaller negative influence on the propensity to step into self-employment for men than for women—in other words, women tend to be much more risk averse than men.

*Professional background:* Recently, Edward Lazear (2002, 2004) proposed a “jack-of-all-trades” theory of entrepreneurship. Based on a coherent model of the choice between self-employment and paid employment, he shows that having a background in a large number of different roles increases the probability of becoming an entrepreneur. The intuition behind this proposition is that entrepreneurs must have sufficient knowledge about a variety of issues to combine the many ingredients needed for survival and success in a business. For paid employees in contrast, it suffices and pays to be a specialist in the field demanded by the job taken. Lazear (2002, 2004) and Wagner (2003a) show that this theory is supported by empirical results for self-employed vs. paid employees in the U.S. and in Germany, respectively. Using the REM data (mentioned above) Wagner (2003b) tests the jack-of-all-trades hypothesis for nascent entrepreneurs vs. persons who decide to continue working as paid employees. He finds evidence of a *ceteris paribus* positive impact of both the number of fields of professional experience and the number of professional degrees for the decision to become a nascent entrepreneur.

*Failure in the past:* Folklore has it that the comparatively low proportion of self-employed in Germany is in part due to a habit that might be termed “stigmatization of failure”: taking a second chance to build one’s own firm after failing as a self-employed person is said to be much more difficult in Germany than in other countries. Wagner (2003c) uses the REM data (mentioned above)

to document that 8% of all people whose former firm went out of business are nascent entrepreneurs today, while the share of failed entrepreneurs among the nascent entrepreneurs is 23%. He investigates the determinants of such a restart. It turns out that both individual and regional factors are important for taking a second chance: this probability is negatively related to age, a high risk aversion and the share of persons in the region who failed in the past, while it is positively related to personal contacts with a young entrepreneur and the regional share of nascent entrepreneurs.

*Regional characteristics:* Two stylized facts emerged from a number of empirical studies for many countries—new venture entry rates differ between regions and the propensity to become an entrepreneur is influenced by socio-demographic variables and attitudes. Wagner and Sternberg (2004) develop a theoretical framework to discuss this link and test whether, for a person of a given age, degree of schooling, attitude toward risk and regional variables and policies matter for the decision to start a new business *ceteris paribus*. Using the REM data (mentioned above) they find that the propensity to be a nascent entrepreneur is higher for people who live in more densely populated and faster growing regions with higher rates of new firm formation, while high prices of land have the opposite impact. Interestingly, it does not matter whether the region has a “left wing” or “right wing” government.

*Characteristics of the (former) workplace:* A stylized fact emerging from a number of empirical studies on the inter-regional differences in new firm formation is that the start-up rate in a region tends to be positively related to the share of employees working in small firms, or the proportion of small firms among all firms in the region (see, e.g., Audretsch and Fritsch, 1994; Gerlach and Wagner, 1994; Reynolds et al., 1994; Armington and Acs, 2002). A similar point has been made in studies dealing with inter-industry differences in new firm formation (see, e.g., Beesley and Hamilton, 1984). A theoretical explanation for this empirical regularity argues that working in a small firm tends to provide employees with a much more relevant experience for starting a new business (e.g., contacts with customers and with the owner of the firm who therefore provides a role model to follow) than working in a large firm (see, e.g., Johnson, 1986 and Mason, 1991). Furthermore, it is well known that job quality tends to be lower in smaller firms along many dimensions, including wages, fringe benefits, job security, participation and opportunities for skill enhancement (see Wagner, 1997 for a survey of the German evidence). Therefore, it is possible that workers are more likely to leave small firms rather than large firms to step into self-employment because of this lower quality of jobs. If this arguments holds, one should expect that people who are working in a small firm (or did so in the past) should have a higher propensity to step into self-employment than others who work(ed) for a large enterprise. A similar argument can be made for those who work(ed) in young firms compared to those

in old firms: Through a close contact with a successful entrepreneur, employees in a young firm have the opportunity to gather information about the transition from paid employment to self-employment with all its problems and about possible solutions. The “employer-as-a-role-model” argument put forward in the context of the small firm should be even more relevant here, because not all small firms are young (and, therefore, not all owners of small firms are role models for potential starters of new firms today), but most of the young firms are small. And we expect it to be most relevant in the case of work experience gathered in young and small firms. Using the REM data (mentioned above) Wagner (2004b) tests the hypothesis that young and small firms are hothouses for nascent entrepreneurs, controlling for various individual characteristics and attitudes. He finds that work experience in a firm that is both young and small is statistically significant and economically important for the decision to become a nascent entrepreneur.

The studies reviewed above that focus on the *ceteris paribus* impact of specific personal characteristics or on selected elements of the environment a person lives and works in on the decision to start creating a new venture shed some light on important aspects of nascent entrepreneurship. However, given that they each are based on a single data set from a single country, collected in a single point in time, it is an open question whether the results are valid in general. Hopefully, further research attempting to replicate these findings using different data sets will tell. And, obviously, there are many aspects related to the determinants of nascent entrepreneurship that are waiting for theoretical and empirical investigations as well.

## 6. WHAT HAPPENS TO NASCENT ENTREPRENEURS AND WHY?

Not all nascent entrepreneurs see their vision through to an eventual start-up in some given period of time (say, in a year after they outed themselves as nascent entrepreneurs in a survey)—some give up and others are still trying. Several studies report empirical findings on the proportions of these sub-groups and on variables that differentiate between them. This literature is surveyed in this section. We summarize the core findings country by country, starting with North America (United States and Canada) and then turning to Europe (Austria, Germany, Italy, the Netherlands, Sweden and Norway) and look at differences and similarities across space afterward.

*United States:* In a pioneering-study, Katz (1990) used data from the Panel Study of Income Dynamics for 1968 to 1972. Of the 2251 wage-or-salaried employees who participated in the survey in 1968, 33 aspired to self-employment. Of these, 27 (or 1.2% of all paid employees) made some effort to prepare themselves for self-employment; these come close to what we call

nascent entrepreneurs today. Of these 27, only six (or 22%) eventually became self-employed between 1968 and 1972. Note that no details are reported in what respect these 6 starters differ from the 19 nonstarters.

Using data for 71 nascent entrepreneurs (taken from two representative samples of 683 adult residents in Wisconsin and of 1016 adult residents of the United States conducted between 1992 and 1993) which were re-interviewed six to eighteen months after their initial interview, Carter et al. (1996) report that between the first and the second interview, 48% of the nascent entrepreneurs had set up a business in operation. 22% had given up and were no longer actively trying to establish a new venture, while 30% report that they were still trying to establish a firm. The authors present what they term "activity profiles" of these three types of nascent entrepreneurs. They suggest that nascent entrepreneurs who were able to start a business were more aggressive in making their business real, acting with a greater level of intensity and undertaking more activities than those people who did not start. Those who gave up performed a pattern of activities that seems to indicate that they discovered that their initial idea for business would not lead to success. Those who are still trying are characterized as not putting enough effort into the start-up process in order to find out whether they should start the business or give up.

Reynolds and White (1997, Chap. 4) use data from the same surveys as Carter et al. (1996), but distinguish four different outcomes (proportions given in brackets): New firm established (45%); actively working on the start-up (28%); temporarily inactive (11%); given up on new business (16%). The authors ask what factors known about the start-up teams and their efforts might differentiate these outcomes and they look at characteristics of the respondent, selected features of the business effort and the activities pursued in starting the business. Important findings include: Men are twice as likely as women to report the business is operating; the proportion of start-ups decreases systematically as educational attainment increases; the proportion of business births is highest for those with intermediate levels of income; most individual attributes, as well as measures of judgment or attitudes, however, have no relationship with the start-up outcome; there are some small effects associated with the economic sector in which the firm operates; and the actual level of effort and investments in the start-up was substantially greater for start-ups that resulted in a firm birth.

Evidently, the samples of U.S. nascent entrepreneurs traced over time in the studies reviewed here are extremely small and the results reported are, therefore, not very reliable. The Panel Study of Entrepreneurial Dynamics (PSED) that involves detailed information on a longitudinal sample of 830 nascent entrepreneurs provides a much better data base for empirical investigations on the topics dealt with in this section. According to Reynolds et al. (2004a, p. 282), however, studies using the PSED longitudinal sample are, to date, primarily at the working paper and conference presentation stages.

*Canada:* Diochon et al. (2003) track the start-up efforts of 151 Canadian nascent entrepreneurs over a two-year period (2000–2002). After 12 (24) months, 29.8% (25.2%) had established an operating business, 33.8% (5.3%) were still trying, 11.2% (5.3%) were inactive, 12.6% (25.2%) had given up entirely and 12.6% (21.1%) could not be reached. Exploring the role individual-level factors play in sustaining efforts to start a business, the authors find no significant differences in personal background factors (socio-demographic, work and career backgrounds), but certain aspects of personal context and personal pre-dispositions are shown to differentiate those who disengaged from the start-up process from those who persevered. It turns out that problem-solving style and goal orientation are especially significant.

*Austria:* Kessler and Frank (2004) analyze data from a longitudinal study in which 290 nascent Austrian entrepreneurs were monitored over a period of three years from 1998 to 2001. At the end, 54.9% of these 1998-nascent had started a business. Those who did not included 7.2% who were still trying and 37.9% who gave up. From a binary logistic regression with “sustained start-up success” as the dependent variable the authors conclude that experience with entrepreneurial thinking, start-ups in the area of crafts and trades and services, full-time business activity, a higher indicated start-up probability at the time of the initial survey and being male are positively related to the probability that a new venture emerges, while those who planned their endeavors jointly with others (team start-ups) were only half as likely to realize their start-ups.

*Germany:* Bahß et al. (2003) use data from the KfW-Gründungsmonitor project to investigate how many of those persons who stated in April–July 2002 that they intend to step into self-employment during the next six month did so until February 2003. From the 300 participants in this follow-up survey, 29% were indeed self-employed, 21% were still trying, 32% delayed their project and 18% gave up. The authors mention that the unemployed more often abandon the process of new venture creation compared with paid employees; and that “starters” and “stoppers” do not differ in important personal characteristics like risk aversion and aspiration for independence (details, however, are not reported). Given that those who state in a survey that they intend to become self-employed in the next half year can not be considered to be nascent entrepreneurs according to the definition given in Section 1, these findings are not strictly comparable with the results reported in other studies reviewed here. However, they provide the only information available for Germany that at least comes close, given that no longitudinal study on German nascent entrepreneurs has yet been performed.

*Italy:* Vivarelli (2004) explores a database including 365 Italian “potential entrepreneurs” who were interviewed in the first quarter of 1999. He considers these individuals to be “potential entrepreneurs” because they

attended—during the ‘1090s—special training courses for people intending to found a new firm. Note that this concept of a potential entrepreneur differs widely from that of a nascent entrepreneur. At the time of the interview, 59% had actually started a new economic venture, while 41% had definitely given up. In a probit equation, the probability of starting is positively and statistically significantly related to the start-up decision being the best choice (opposed to more defensive motives), a high level of information, no free admittance to the training course and writing a business plan.

*Netherlands:* van Gelderen et al. (2001) followed 330 nascent entrepreneurs identified in the fall of 1998 over a one year period and asked for the current status of the start-up effort. 47% started their business, 27% were still organizing and 26% had abandoned the effort. They report that in comparison to people who gave up, starters are entrepreneurs already, have more industry experience, start out with less start-up capital, use fewer third-party loans and start out in manufacturing. Compared to those who are still organizing, starters are relatively often male, entrepreneur and want to start full-time. In a follow-up study, van Gelderen et al. (2003) report that after three years, a minimum of 36% of the sample started and a minimum of 20% abandoned the start-up effort, while there is no information about the eventual start-up status of the remaining 44%. A comparison of those who succeed in starting a business and those who abandon the start-up effort reveals that significant variables include start-up capital (nascents who intend to use more start-up capital have lower probabilities to get their business running) and perceived risk of the market, starting a manufacturing firm and starting full time. None of the included individual characteristics seem to distinguish successful nascent entrepreneurs from the unsuccessful ones.

*Sweden:* Davidsson and Honig (2003) followed 380 Swedish nascent entrepreneurs first interviewed between May and September 1998 for 18 months. They use the occurrence of a first sale during these 18 months as an indication of a nascent firm’s eventual emergence. Sixty two percent of the nascent entrepreneurs reported first sales during this period. In a logistic regression, the probability of having a first sale turns out to be unrelated to several measures of human capital (years of education, business class taken, years experience as manager, years work experience and previous start-up experience) and to age and gender. Among the social capital variables, only being member of a business network and having close friends or neighbors in business have a statistically significant positive impact on the probability of a first sale.

*Norway:* Alsos and Ljunggren (1998) report that from 149 Norwegian nascent entrepreneurs interviewed first in a survey conducted early in 1996, 46% started a business when re-interviewed 12 months later, 25% were still trying and 29% gave up. These proportions were identical for men and women.

Some but not all of the studies reviewed in this section follow, explicitly or implicitly, but sometimes only partly, the research design of the Panel Study of Entrepreneurial Dynamics (PSED) discussed in Reynolds (2000). Comparability across space, therefore, is limited. Furthermore, the rather small and sometimes tiny samples, different time frames for follow-up studies and different specifications of the empirical models used make it impossible to draw any definite conclusions. However, at least two tentative conclusions emerge: First, a significant fraction of nascent entrepreneurs—between one in two and one in three—step into the next phase, becoming infant entrepreneurs in the year following the first survey. Second, observed individual characteristics tend to play a minor role only in differentiating between who starts and who gives up.

## 7. CONCLUDING REMARKS

While we knew next to nothing about nascent entrepreneurs ten years ago, thanks to the joint effort of a group of researchers, most of whom are affiliated with the Global Entrepreneurship Monitor (GEM) project, we now have reliable information on the share of nascent entrepreneurs in the population of a large number of countries; the reason for differences in this share across space and time, however, is less well understood. Furthermore, we have a sound knowledge about the prevalence of nascent entrepreneurs in certain sub-groups (like males and females, or people with various educational backgrounds). Less is known about precisely what nascent entrepreneurs are doing and about the timing of the activities. The same conclusion holds with respect to factors that are important for becoming a nascent entrepreneur and for crossing the threshold between nascent and infant entrepreneurship: But a lack of comparability among the numerous empirical studies for different countries makes it impossible to draw any definite conclusions.

Stylized facts that could be most valuable for entrepreneurship researchers, policy makers and, last but not least, nascent entrepreneurs, need to be based on results from a number of studies using large, comprehensive longitudinal data bases that are comparable across countries and that can be accessed by researchers for replication and extension of former studies. The Global Entrepreneurship Monitor (GEM) and the Panel Study of Entrepreneurial Dynamics (PSED) projects and the data collected within these projects, are important steps toward this aim. The importance of new firms for economic dynamics and of nascent entrepreneurs for new firms, points to the need for further steps in the future.



## NOTE

<sup>1</sup> Financial support by the German Research Foundation DFG under grants WA 610/2-1 and WA 610/2-2 for the project *Regional Entrepreneurship Monitor (REM) Germany* is gratefully acknowledged. I thank Heiko Bergmann, Simon Parker, Claus Schnabel, Friederike Welter and an anonymous referee for helpful comments on an earlier version. The usual disclaimer applies.

## REFERENCES

- Acs, Z. J., Arenius, P., Hay, M. and Minniti, M. (2005). *Global Entrepreneurship Monitor 2004 Executive Report*. Babson College and London Business School.
- Alsos, G. A. and Ljunggren, E. (1998). Does the Business Start-Up Process Differ by Gender? A Longitudinal Study of Nascent Entrepreneurs. In *Frontiers of Entrepreneurship Research*. Babson College, Wellesley, MA.
- Armington, C. and Acs, Z. J. (2002). The Determinants of Regional Variation in New Firm Formation. *Regional Studies* 36: 33–45.
- Audretsch, D. B. and Fritsch, M. (1994). The Geography of Firm Births in Germany. *Regional Studies* 28: 359–365.
- Bahß, C., Lehnert, N. and Reents, N. (2003). Warum manche Gründungen nicht zustande kommen. *KfW-Research*, Nr. 10, Oktober, 2–9.
- Beesley, M. E. and Hamilton, R. T. (1984). Small Firms' Seedbed Role and the Concept of Turbulence. *Journal of Industrial Economics* 33: 217–231.
- Bergmann, H., Japsen, A. and Tamas, C. (2002). *Regionaler Entrepreneurship Monitor (REM). Gründungsaktivitäten und Rahmenbedingungen in zehn deutschen Regionen*. Universität zu Köln, Universität Lüneburg.
- Blanchflower, D. G. (2004). Self-employment: More may not be Better. *Swedish Economic Policy Review* 11 (2): 15–74.
- Blanchflower, D. G., Oswald, A. and Stutzer, A. (2001). Latent Entrepreneurship across Nations. *European Economic Review* 45: 680–691.
- Carree, M. and Thurik, R. (2005). *Understanding the Role of Entrepreneurship for Economic Growth*. Discussion Papers on Entrepreneurship, Growth and Public Policy, MPI Jena, #1005, February.
- Carter, N. M. and Kolvereid, L. (1998). Women Starting New Businesses: The Experience in Norway and the United States. In *Women Entrepreneurs in Small and Medium Enterprises*, pp. 185–202. OECD, Paris.
- Carter, N. M., Gartner, W. B. and Reynolds, P. D. (1996). Exploring Start-up Event Sequences. *Journal of Business Venturing* 11: 151–166.
- Davidsson, P. (2004). *Researching Entrepreneurship*. Springer, Boston.
- Davidsson, P. and Honig, B. (2003). The Role of Social and Human Capital among Nascent Entrepreneurs. *Journal of Business Venturing* 18: 301–331.
- Delmar, F. and Davidsson, P. (2000). Where do they Come From? Prevalence and Characteristics of Nascent Entrepreneurs. *Entrepreneurship & Regional Development* 12: 1–23.
- Diochon, M., Menzies, T. V. and Gasse, Y. (2003). *Insight into the Dynamics of Canadian Nascent Entrepreneurs' Start-up Effort and the Role Individual Factors Play in the Process*. Mimeo, St. Francis Xavier University, Brock University and Université Laval.
- Diochon, M., Gasse, Y., Menzies, T. and Garand, D. (2001). From Conception to Inception: Initial findings from the Canadian Study on Entrepreneurial Emergence. *Proceedings of the Administrative Science Association of Canada, Entrepreneurship Division* 22 (21): 41–51.

- Gartner, W. B. and Carter, N. M. (2003). Entrepreneurial Behavior and Firm Organizing Processes. In *Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds), pp. 195–221. Kluwer Academic Publishers, Boston, MA.
- Gartner, W. B., Shaver, K. G., Carter, N. M. and Reynolds, P. D. (Eds) (2004). *Handbook of Entrepreneurial Dynamics. The Process of Business Creation*. Sage Publications, Thousand Oaks.
- van Gelderen, M., Bosma, N. and Thurik, R. (2001). *Setting up a Business in the Netherlands: Who Starts, Who Gives up, Who Is Still Trying*. Paper presented at the Babson 2001 Entrepreneurship Conference.
- van Gelderen, M., Thurik, R. and Bosma, N. (2003). *Success and Risk Factors in the Pre-start-up Phase*. Mimeo, Vrije Universiteit Amsterdam, September.
- Gerlach, K. and Wagner, J. (1994). Regional Differences in Small Firm Entry in Manufacturing Industries: Lower Saxony, 1979–1991. *Entrepreneurship and Regional Development* 6: 63–80.
- Johnson, P. S. (1986). *New Firms: An Economic Perspective*. Allen and Unwin, London.
- Katz, J. A. (1990). Longitudinal Analysis of Self-employment Follow-through. *Entrepreneurship & Regional Development* 2: 15–25.
- Kessler, A. and Frank, H. (2004). *Nascent Entrepreneurship in Longitudinal Perspective—The Impact of Personality, Environment, Resources and the Founding Process on the Decision to Start Business Activity: Results from the Vienna Entrepreneurship Studies*. Paper presented at Babson Kauffman Entrepreneurship Research Conference (BKERC) 2004, June 3–5, University of Strathclyde, Glasgow, Scotland.
- Kim, P. H., Aldrich, H. E. and Keister, L. A. (2003). *If I Were Rich? The Impact of Financial and Human Capital on Becoming a Nascent Entrepreneur*. University of North Carolina at Chapel Hill and Ohio State University, draft mimeo, January.
- Lazear, E. P. (2002). *Entrepreneurship*. Working Paper #9109, National Bureau of Economic Research, Cambridge, MA.
- Lazear, E. P. (2004). Balanced Skills and Entrepreneurship. *American Economic Review—Papers and Proceedings* 94 (2): 208–211.
- Lückgen, I. and Oberschachtsiek, D. (2004). *Regionaler Entrepreneurship Monitor (REM II 2003/2004). Zur Dynamik von Gründungsaktivitäten in Deutschland: Ausmaß und Ursachen*. Universität zu Köln, Universität Lüneburg.
- Mason, C. (1991). Spatial Variations in Enterprise: The Geography of New Firm Formation. In *Deciphering the Enterprise Culture. Entrepreneurship, Petty Capitalism and the Restructuring of Britain*, R. Burrows (Ed.). Routledge, London and New York.
- Minniti, M. and Arenius, P. (2003). *Women in Entrepreneurship*. Paper presented at The Entrepreneurial Advantage of Nations: First Annual Global Entrepreneurship Symposium, United Nations Headquarters, April 29, 2003.
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Porter, M. E. and Stern, S. (2002). National Innovative Capacity. In *The Global Competitiveness Report 2001-2002*, M. E. Porter et al. (Eds), pp. 102–118. Oxford University Press, New York.
- Reynolds, P. D. (1997). Who Starts New Firms?—Preliminary Explorations of Firms-in-Gestation. *Small Business Economics* 9: 449–462.
- Reynolds, P. D. (2000). National Panel Study of U.S. Business Start-ups: Background and Methodology. In *Databases for the Study of Entrepreneurship*, J. A. Katz (Ed.), pp. 153–227. JAI Press, Amsterdam.
- Reynolds, P. D. and White, S. B. (1997). *The Entrepreneurial Process. Economic Growth, Men, Women and Minorities*. Quorum Books, Westport, Connecticut and London.
- Reynolds, P. D., Storey, D. and Westhead, P. (1994). Regional Variations in New Firm Formation. *Regional Studies* 28: 443–456.

- Reynolds, P. D., Hay, M. and Camp, S. M. (1999). *Global Entrepreneurship Monitor 1999 Executive Report*. Kauffman Center for Entrepreneurial Leadership at the Ewing Kaufman Foundation.
- Reynolds, P. D., Hay, M., Bygrave, W. D., Camp, S. M. and Autio, E. (2000). *Global Entrepreneurship Monitor 2000 Executive Report*. Kauffman Center for Entrepreneurial Leadership at the Ewing Kaufman Foundation.
- Reynolds, P. D., Camp, S. M., Bygrave, W. D., Autio, E. and Hay, M. (2001). *Global Entrepreneurship Monitor 2001 Summary Report*. London Business School and Babson College.
- Reynolds, P. D., Bygrave, W. D., Autio, E., Cox, L. W. and Hay, M. (2002a). *Global Entrepreneurship Monitor 2002 Executive Report*. Babson College, Ewing Marion Kauffman Foundation and London Business School.
- Reynolds, P. D., Carter, N. M., Gartner, W. B., Greene, P. G. and Cox, L. W. (2002b). *The Entrepreneur Next Door. Characteristics of Individuals Starting Companies in America. An Executive Summary of the Panel Study of Entrepreneurial Dynamics*. Ewing Marion Kauffman Foundation, Kansas City, Missouri.
- Reynolds, P. D., Carter, N. M., Gartner, W. B. and Greene, P. G. (2004a). The Prevalence of Nascent Entrepreneurs in the United States: Evidence from the Panel Study of Entrepreneurial Dynamics. *Small Business Economics* 23: 263–284.
- Reynolds, P. D., Bygrave, W. D. and Autio, E. (2004b). *Global Entrepreneurship Monitor 2003 Executive Report*. Babson College, London Business School and Kauffman Foundation.
- Shaver, K. G., Carter, N. M., Gartner, W. B. and Reynolds, P. D. (2001). *Who is a Nascent Entrepreneur? Decision Rules for Identifying and Selecting Entrepreneurs in the Panel Study of Entrepreneurial Dynamics (PSED)*. Paper presented at the Babson College Kauffman Foundation Entrepreneurship Research Conference, Jönköping, Sweden.
- van Stel, A., Wennekers, S., Thurik, R., Reynolds, P. D. and de Wit, G. (2003). *Explaining Nascent Entrepreneurship across Countries*. SCALES-paper N200301, EIM Business and Policy Research, May.
- Sternberg, R. (2000). *Entrepreneurship in Deutschland. Das Gründungsgeschehen im internationalen Vergleich*. Edition Sigma, Berlin.
- Uhlaner, L. and Thurik, A. R. (2004). *Post-Materialism: A Cultural Factor Influencing Total Entrepreneurial Activity Across Nations*. Mimeo, Erasmus University Rotterdam, February.
- Verheul, I. and Thurik, A. R. (2003). *Explaining the Entrepreneurial Activity Rate of Women: A Macro-level Perspective*. SCALES Paper N200304, EIM Business and Policy Research.
- Vivarelli, M. (2004). Are All the Potential Entrepreneurs so Good? *Small Business Economics* 23: 41–49.
- Wagner, J. (1997). Firm Size and Job Quality: A Survey of the Evidence from Germany. *Small Business Economics* 9: 411–425.
- Wagner, J. (2003a). Testing Lazear's Jack-of-All-Trades View of Entrepreneurship with German Micro Data. *Applied Economics Letters* 10: 687–89.
- Wagner, J. (2003b). *Are Nascent Entrepreneurs Jacks-of-All-Trades? A Test of Lazear's Theory of Entrepreneurship with German Data*. Institute for the Study of Labor IZA Discussion Paper No. 911 (forthcoming, *Applied Economics*).
- Wagner, J. (2003c). Taking a Second Chance: Entrepreneurial Re-starters in Germany. *Applied Economics Quarterly* 49: 255–272.
- Wagner, J. (2004a). *What a Difference a Y Makes: Female and Male Nascent Entrepreneurs in Germany*. Discussion Paper No. 1134, Institute for the Study of Labor, IZA, Bonn, Germany (forthcoming, *Small Business Economics*).
- Wagner, J. (2004b). Are Young and Small Firms Hothouses for Nascent Entrepreneurs? Evidence from German Micro Data. *Applied Economics Quarterly* 50: 379–391.

- Wagner, J. and Sternberg, R. (2004). Start-up Activities, Individual Characteristics and the Regional Milieu: Lessons for Entrepreneurship Support Policies from German Micro Data. *Annals of Regional Science* 38: 219–240.
- Welter, F. (2001). *Nascent Entrepreneurship in Germany*. RWI, Essen.

SHAKER A. ZAHRA

*Center for Entrepreneurial Studies & Department of Strategic Management & Organization,  
Room 3-428, Carlson School of Management, University of Minnesota, 321 19th Ave. South,  
Minneapolis MN 55455, USA  
E-mail: szahra@csom.umn.edu*

### 3. New Venture Strategies: Transforming Caterpillars into Butterflies

#### 1. INTRODUCTION, OBJECTIVES AND SCOPE<sup>1</sup>

Timmons (2005) defines entrepreneurship as “transforming caterpillars into butterflies.” This transformation is most vividly noted in the launch of the new venture, a milestone in the firm’s life (Aldrich, 2000; Carter, Gartner and Reynolds, 1996). By then, an entrepreneur would have searched for and identified a viable opportunity, evaluated it, explored ways to exploit and give it form, chosen a business model, assembled the key resources necessary to bring it to life, selected the company’s “organizational” and legal form, chosen a top management team (TMT) and connected the young firm to its key stakeholders in order to build its name recognition, credibility and market legitimacy. These pre-launch activities set the stage for transforming the business idea into a viable business enterprise (Carter et al., 1996). A new venture needs an effective strategy for the idea to fly, establish its market presence, gain legitimacy and achieve financial success. Like the wings of a butterfly, a strategy helps the new venture to take off.

This chapter reviews the research on new venture strategy and its implications for organizational survival, financial performance and growth. A new venture is a firm that has been in existence for a period of eight years or less. The chapter focuses on the competitive, cooperative and political strategies that entrepreneurs might use following their venture’s launch. Initially, the discussion will focus on the nature of competitive strategy, its key antecedents and effects. The discussion will also cover the cooperative and political strategies that new ventures use, highlighting the importance of

synchronizing the competitive, cooperative and political facets of new venture strategy. The discussion will then recognize key contingencies that influence the relationships among these strategies.

Much has been written about new venture strategy (Bahrami and Evans, 1989; Covin, Slevin and Covin, 1990), especially its competitive dimensions that help the firm to select a chosen niche, define its rivals, select its strategic priorities and leverage its resources in positioning itself and achieving a competitive advantage (Carter, Stearns, Reynolds and Miller, 1994; McDougall, Robinson and Herron, 1994; Romanelli, 1989; Park and Bae, 2004; Sandberg, 1986; Shepherd and Shanley, 1998; Vesper, 1990). These studies have been grounded in the industrial organizational economics (Bain, 1959; Porter, 1980), the resource-based (Barney, 1991, 2001; Barney, Wright and Ketchen, 2001; Penrose, 1959) and the dynamic capabilities (Teece, Pisano and Shuen, 1997; Winter, 2003) views. Each perspective holds different views of the markets, competition and the nature and durability of a firm's competitive advantage. Thus, there are several rich empirical studies that speak to these diverse issues, enriching our understanding of the complex nature of the strategic choices that entrepreneurs make as they enter their markets. Not surprisingly, perhaps, findings of prior results have been inconclusive and sometimes contradictory (Carter et al., 1994; Chandler and Hanks, 1994; Covin et al., 1990). The review I present in this chapter will cover key themes in these divergent views, highlighting their contributions, shortcomings and implications for future research. The chapter will also identify areas where convergence among these views might be taking place, shaping our thinking and theorizing about new venture strategy.

There are two additional and critical dimensions to new venture strategy: cooperative and political. Both dimensions have received limited but growing attention in the literature. Cooperative strategies focus on the conditions under which new ventures should collaborate with other ventures or established companies, rather than going it alone. A significant amount of research has been conducted on the various cooperative strategies that new ventures can use to position themselves, survive and make a profit (McGee, Dowling and Megginson, 1995; Vesper, 1990). Entrepreneurs use political strategies to influence key stakeholders and gain access to the resources they need to survive and thrive. A young but growing body of research has examined these political dimensions of new venture strategy (Alvord, Brown and Letts, 2004; Harding, 2004). This research reveals that not all entrepreneurs are interested in maximizing their financial gains or creating wealth for themselves or their families. Some entrepreneurs pursue other multiple goals such as improving the quality of life in their society and being good citizens who add value to their communities.

Researchers who work on cooperative and political strategies hold very different assumptions about the nature of the firm and its markets from those scholars who study competitive strategy. As our discussion unfolds, however,

it will become evident that these three views (competitive, cooperative and political) are complementary, not rival views of successful new venture strategy. Astute entrepreneurs rapidly appreciate the value each of these views can contribute to their ventures' survival and successful performance.

## 2. COMPETITIVE STRATEGY

Upon launching a new venture, the entrepreneur has to address four related questions: (a) Where will the venture compete (defining the "served market")? (b) What is the appropriate scope of operations (defining the "business scope")? (c) What should the "strategic posture" of the venture be? And (d) what are the key competitive weapons the venture should use? Answers to these related questions form the content of a firm's competitive strategy, defined as its set of coherent choices about its market, scope of operations and how it will compete. Figure 3-1 outlines the various steps that an entrepreneur might undergo to address these issues. Even though Figure 3-1 portrays a sequential series of activities for analytical clarity, these steps are often iterative in nature, with possible multiple loops. Clarity about some of the questions or issues encountered in one step could compel the entrepreneur to reconsider some of the choices or plans conceived in prior stages.

### 2.1. Defining the Served Market

As Figure 3-1 suggests, a key decision the entrepreneur (or the TMT) will have to make is: Which part of the industry or market should the venture target? Data collected about the market at the opportunity analysis stage (Fliesher and Bensoussan, 2003) could guide the entrepreneur's decision making process. When the venture is launched, greater precision is necessary to know who the customers are, how many there are, what they need and how to best serve and approach them. Answering these questions makes it possible for the TMT to determine the venture's *served market*, the part of the market in which they will operate. Given that many new ventures enter foreign markets early in their life cycles (Andersson, 2004), the served market could be further defined domestically or internationally (Zahra, 2005; Zahra and George, 2000b).

Defining the served market determines the ventures' current and future competitors, guides future decisions about the scope of domestic and international operations and provides a foundation for setting realistic goals regarding profitability and growth (Fliesher and Bensoussan, 2003). Yet, delineating the served market is an iterative process that unfolds over time as the TMT gains experience with the industry and its customers. The growing globalization of markets and industries might also cause companies to alter their definition

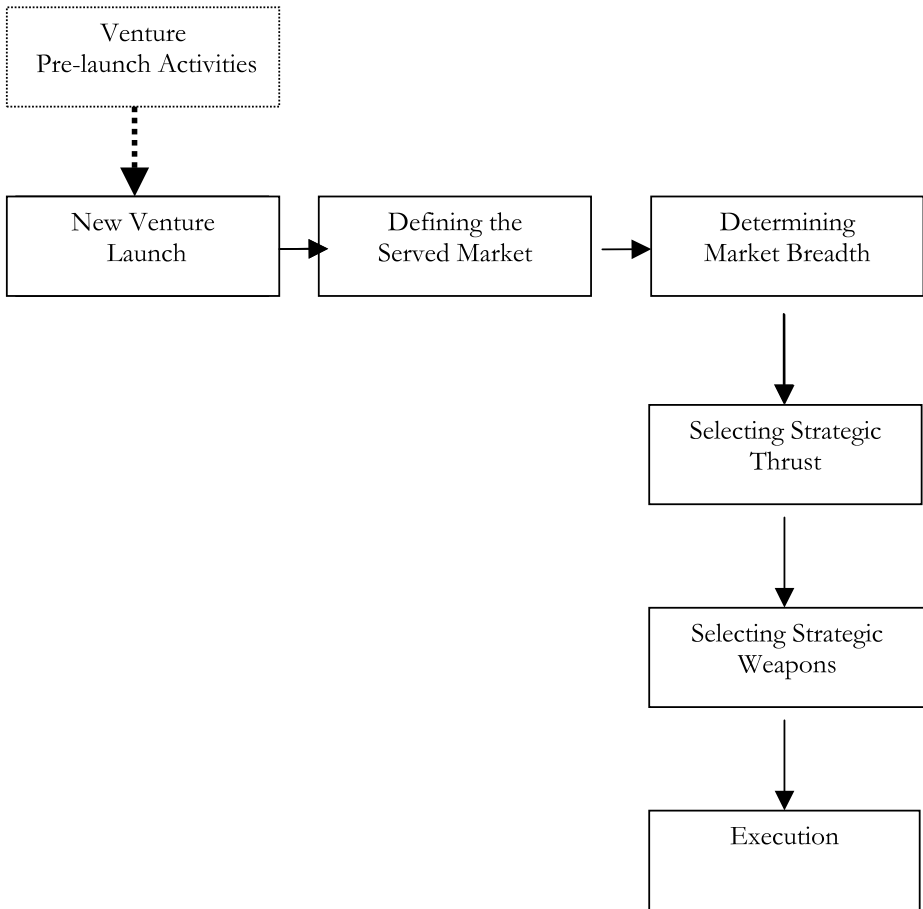


FIGURE 3-1 *Steps followed in mapping out new venture strategy.*

of the served markets. Changes in technology, competitors' composition and strategies, feedback from customers and suppliers, and the TMT experience can shape the venture's redefinition of the served market (Bhide, 2000; Dean and Meyer, 1996; Eisenhardt and Schoonhoven, 1990; Sandberg, 1986; Vesper, 1990).

### 2.2. *Determining the Scope of Operations*

Figure 3-1 indicates that the next question the TMT has to address is: Within the chosen served market, what would be an appropriate scope for the venture? Entrepreneurs seeking guidance from the literature are likely to get contradictory advice. Some researchers stress the need to start small, build the venture's operations and broaden its scope over time. This sequential



approach allows the venture to focus its energy, build a market presence and gain market credibility. New ventures often suffer from “liabilities of newness” (Stinchcombe, 1965), where they are not easily accepted by customers or other companies, limiting their access to capital and other vital resources. The “build as you go” approach also imposes strategic discipline on the TMT, ensuring the pursuit of strategic options that new ventures can afford. It also ensures coherence in the strategies the TMT develops, meaning that they make consistent choices in pursuit of their central goals.

Other researchers suggest that aggressive entry into the market is more likely to be effective (for a discussion, McDougall et al., 1994). They note that new ventures should target a relatively broad market and enter on a large scale, offering multiple products or versions of the same products. This signals a new venture’s strong capabilities, helping to quickly overcome some of the liabilities of newness. Of course, broad market entry is an expensive strategy that only very new ventures can afford. Serving diverse market segments could also be difficult to manage because customers might have different needs and buying habits. Coordinating new ventures’ production, marketing and distribution activities could be challenging for the TMT. Many ventures also specialize in particular technologies and products and cannot always serve multiple segments adequately.

Choices by new ventures are sometimes complicated by the fact that strategic breadth has multiple dimensions (McDougall et al., 1994). The best known dimension is the *number of markets or segments* targeted by the venture in its domestic or international markets. Another key dimension is the *number of products* the venture offers. While related to the number of market segments, a new venture could target the same segment with multiple products or use the same product to reach multiple segments. A final dimension of breadth is the *variety within the same product*; some ventures offer one model but others might develop and introduce several models. Developing capabilities along these various lines could be a time consuming, expensive and challenging process. As a result, some TMTs might select and develop a combination of these dimensions.

New ventures are likely to vary considerably in their strategic breadth because of different definitions of their served market (McDougall et al., 1994). The nature of the industry can also induce variability. For example, growing industries usually have more viable niches for new ventures to target but declining industries have fewer niches (Porter, 1980). The firm’s resources could also shape the definition of its breadth, with well endowed ventures targeting more segments. The profitability and growth targets established by the TMT also play a key role in making this choice. Those teams that seek fast growth might adopt a broader definition of the industry and target multiple segments. Of course, it takes a lot more to transform these goals into reality; big

market shares do not automatically generate higher profitability or ascendance to industry leadership. Credible commitments of resources, developing effective distribution channels, aggressive advertising, building and cultivating synergy across product lines and segments and cleverness in deploying the firm's assets are essential in order to gain an advantage from a broadly defined market.

Several factors are likely to influence entrepreneurs' choice of their firms' market scope. Notably, there are major risks with each choice. Narrow market scope could limit the firm's initial growth rates. Broad market scope, however, could lead to a lack of focus and unfulfilled demand. Each of these options has cost implications as well. For example, broad market scope could require significant resources that the new venture does not have or cannot easily acquire. The desirability of a given market scope depends also on a firm's industry conditions; growing industries might encourage entrepreneurs to broaden their firm's market definition and reap the benefits associated with growth. The nature of the venture's products and the depth and breadth of its product line could also influence the choice of a company's market scope. A venture that has a single or a few products is better off targeting a narrow market or a few selected segments. Conversely, if the venture has multiple product lines, it might opt to define its market broadly.

The objectives and goals of founders and other TMT members can also determine the definition of a company's market scope. When managers are eager to achieve fast growth and create wealth, they might pursue a broad market. Entrepreneurs' personalities might also influence the choice of the market scope as well. Some entrepreneurs are adventuresome and are apt to have a more optimistic view of the industry and its growth trajectories. These entrepreneurs might decide to pursue a broadly defined market.

### 2.3. *Choosing the Strategic Thrust*

Once the scope of the venture's operations has been chosen, the next step is selecting its strategic thrust: *how it will compete* (Figure 3-1). This step requires careful attention to several related and iterative issues. First, the entrepreneur has to decide if the new venture intends to compete as "first to the market" or "second to the market" or simply adopt a "me-too" strategy. Each of these choices has pluses and minuses and several factors come to bear on making this decision.

A *first to the market* strategy gives the new venture an opportunity to define its competitive arena in a way that fosters its goals, especially growth. According to this strategy, the venture aims to be among the pioneers in developing and introducing innovative products, goods or services (Porter, 1985). To do so, the venture has to focus on radical innovation, retain first

class manufacturing, control strong distribution channels and use aggressive advertising and promotion. This is a costly strategy and some ventures have cleverly used strategic alliances and reduced their investments and aggressively competed as pioneers. Market pioneers are the first to conceive of an industry's boundaries and work hard to create and develop that industry. Technological pioneers are companies that lead the development of radically new products in their industries (Zahra, Nash and Bickford, 1995). Technological pioneers might evolve into market or industry pioneers but some remain effective and successful niche players. One advantage of this strategy is that the firm can skim the market by targeting lucrative segments. Obviously, this strategy could be used with a particular technology to develop a unique niche or could be applied to multiple segments. Besides the significant financial resources needed, industry and technological pioneers often fail to retain their positions for very long (Grant, 1998; Porter, 1980; Zahra et al., 1995). Ironically, the success of the pioneers often attracts well-established companies to enter the market, either by creating their own units to compete with independent ventures or by acquiring these successful ventures. Over time, some pioneers become inert in their decisions; others might fall behind new entrants who bring new technologies, marketing and distribution skills or business models and leapfrog existing ventures.

A *second to the market* strategic thrust means following the industry leaders by quickly imitating what they are doing and offering viable substitutes (Porter, 1985). This requires excellence in applied engineering and R&D. Strong marketing and distribution skills also allow new ventures to offer new versions of the same products and differentiate what they are doing from their competition. Consequently, this strategy is viable in those industries where well-healed established industry and technological leaders enjoy strong name recognition and a relatively large market share. Some clever entrepreneurs have followed this strategy to enter the industry, without challenging established companies, and progressively build new skills that have allowed them to leapfrog these leaders.

A *"me-too"* strategic thrust emphasizes imitation combined with a lack of product differentiation. A venture might use this strategy to reduce its investments in R&D and marketing; avoiding head-to-head competition with established companies. The new venture might focus more on uncovering relatively smaller market niches that have been overlooked by existing companies. These ventures charge lower prices or offer specialized or personal services that build relationships with customers. Continuous, incremental processes or product innovations aid these ventures in sustaining and defending their market positions.

## 2.4. Choosing the Strategic Weapons

As Figure 3-1 indicates, selecting the strategic thrust sets the foundation for making decisions about the product, price, promotion and distribution activities and effectively position the new firm in its market (Covin et al., 1990; Li and Atuahene-Gima, 2001; McDougall et al., 1994). These decisions should not and can not be made in isolation from the firm's external environment or from each other (Porter, 1980; 1985). The entrepreneur should capitalize on the interrelatedness of these choices to offer a unique bundle of utility for the customer. In particular, the entrepreneur has to address the need for strategic coherence while ensuring flexibility, thus maximizing the degrees of freedom the firm will have as in building its market presence.

Researchers have proposed typologies that depict the various choices associated with the strategic thrusts new ventures might use. The Miles and Snow (1978) and Porter (1980) typologies are especially noteworthy. These typologies are an important means of identifying different strategies and the skills required for each. Even though these typologies are simplistic in that they emphasize extreme "types," they highlight the consistent choices that companies have to make in pursuit of each.

*2.4.1. The Miles and Snow Typology* Miles and Snow (1978) offer a widely recognized typology of the strategic thrusts that companies might use. This typology has inspired a significant amount of research and debate, providing support of its validity and key propositions (for an extensive review, see Zahra and Pearce, 1990). Briefly, Miles and Snow (1978) argue that the TMT has to address three interrelated problems that influence the success of a company: *entrepreneurial* (what is the domain of the business), *engineering* (how will the business create goods and services that actualize the mission of the firm) and *administrative* (how will the business create the appropriate organizational form). Answers to these three "problems" highlight three viable strategic types: Defenders, Prospectors and Analyzers. Miles and Snow also identify a dysfunctional type: Reactors.

*Defenders* are specialized ventures that do well in relatively stable business environments. They have a narrow market scope, thrive on efficient operations and compete primarily based on low price. These ventures have effective manufacturing bases and highly integrated distribution channels. Defenders succeed based on efficiency and therefore have to aggressively pursue process innovation. Many new ventures that are launched in maturing or declining industries follow the Defender strategy for years.

*Prospectors* usually thrive in dynamic markets. They constantly change their business definitions based on their recognition of the new opportunities that become open as their industries change. Given the dynamism of their environments, flexibility through innovation is a strategic priority for Prospec-

tors who have to constantly introduce new products to serve new markets or segments. Prospectors spend more heavily on R&D, marketing research, advertising and distribution than Defenders. New ventures that are launched in emerging or growing industries are likely to adopt the Prospector strategy.

Miles and Snow view competitive strategies along a continuum, ranging from Defenders to Prospectors with Analyzers placed in the middle. *Analyzer* new ventures combine the attributes of both Defenders and Prospectors. They have a stable core business but pursue new opportunities outside that core. Analyzers use the Defender strategy in their core business and the Prospector strategy in their new fields of business. For many newly launched ventures, managing the duality of these strategies can be challenging. Having an organizational structure that handles this duality could be especially taxing.

Finally, there are the *Reactor* organizations which seem unable or unwilling to commit themselves or their resources in a coherent fashion. They change their business definition frequently and move from one strategy to another. This lack of discipline often reflects a pathological and indecisive leadership style inconsistent with long-term, strategic commitment. Reactors are more likely to fail than Defenders, Analyzers or Prospectors. Their poor leadership and unstable choice of strategies are primary causes of their failure.

**2.4.2. Porter's Typology** Porter (1980, 1985) has also proposed a widely used typology that highlights two dimensions of strategic choices: Scope (narrow vs. broad) and Quality (low vs. high). When these options are combined, several types of strategies become evident. Some new ventures might use a *focus strategy* by targeting narrowly defined markets. When these ventures offer low quality and charge low prices, they follow the Defender strategy, discussed above. If new ventures target a broadly defined market, offering low quality and charging low prices, they follow *the low cost strategy* outlined earlier. New ventures that value innovation, high quality and charging premium prices follow a "*differentiation*" strategy. This strategy could be used also in a broadly or narrowly defined market. Innovative technologies have made it feasible for these ventures to introduce high quality and innovative products without charging higher prices, giving them a competitive advantage.

Regardless of the strategic thrust used, entrepreneurs succeed by: (a) making sure that the strategic weapons chosen are consistent with that thrust; (b) ensuring that the strategic weapons support and reinforce each other by creating synergy; and (c) maintaining the strategic flexibility necessary to overcome unexpected changes in market forces or competitive dynamics. Strategic flexibility stems from having resources that could be used for different purposes (Sanchez, 1995). It also results from the creativity of the TMT in finding new ways to reconfigure these resources and build a competitive advantage. These issues influence the selection of new ventures' TMT. This

team should be diverse in its skills and decision making styles in order to ensure creativity. But this creativity could be a double-edge sword. On one hand, it expedites the recognition of new strategic alternatives. But, on the other hand, it could induce needless or dysfunctional strategic change that could be as dangerous as conservatism. The diversity of the TMT may lead also to cognitive (disagreements on ends and means) and affective conflicts (triggered by feelings and attitudes toward issues of interest). Persistent dysfunctional conflicts could paralyze the new ventures and slow the pace of their evolution.

## *2.5. Determining the Venture' Competitive Advantage*

Newly launched ventures that develop and execute an effective strategy are likely to gain a competitive advantage over their rivals. Competitive advantage has been defined differently in the literature (Porter, 1985), but most definitions converge on the ability of the firm to gain and sustain higher performance as compared to its rivals (Grant, 1998). Figure 3-1 portrays the sequence of activities and the corresponding decisions that entrepreneurs take to position the firm to gain market share, make profits and even grow. The major assumption underlying these decisions is that effective positioning leads to a competitive advantage. However, as competition intensifies, ventures often imitate each other and over time might lose their uniqueness. Researchers have investigated some of the key sources of enduring competitive advantage. Research findings highlight the role of organizational resources and capabilities in this regard.

*2.5.1. Resources and Competitive Advantage* For years, strategy researchers have sought to identify the variables that give a firm a differential advantage over its rivals. This advantage frequently lies in the way the firm is positioned to acquire market share and then use this market to build an enduring source of advantage (for a review, see Grant, 1998; Shepherd and Shanley, 1998; Vesper, 1990). Positioning centers on selecting viable market niches, targeting the right segments and deploying resources effectively.

Some scholars have charged that this “positioning” perspective on strategy reduces competition into a series of jockeying for positions as an industry continues to evolve. Over time, some companies lose their uniqueness, making it easier for their rivals to steal away market share. Consequently, some have suggested that the sources of distinctiveness might lie in a firm’s resources and decision-making processes. They observe that some resources are more strategically valuable than others and therefore can shape the strategic choices companies make (Penrose, 1959). Barney (1991, 2001a, 2001b), in particular, has argued that resources that are valuable, rare, inimitable and organizationally embedded can give new ventures competitive advantages over their rivals. Other

researchers have also noted that it is not only the *stock* of these resources that makes the difference but also *the flow* of certain resources that can make the difference in new ventures' long-term performance (for a review, see DeCarolis and Deeds, 1999; Deeds et al., 1999).

Researchers have invoked the resource-based view (RBV) to propose that new ventures can leverage their resources in ways that allow them to gain and protect their advantages over time. They have also asserted that the unique managerial decision-making processes that exist within new ventures can give them an advantage over their established competitors. For example, founders' knowledge of the industry and control of their firms' operations and the psychological stake these founders have in the well-being and success of their ventures can determine their strategic moves. Entrepreneurial qualities (e.g., aggressiveness, risk-taking) can also result in a strategy that differs markedly from those of established companies. It is this intimate link among founders, the firm and resources that can influence their strategic choices and any changes they initiate in the new ventures' strategies. Findings of recent research revealed on the influence on entrepreneurial intentions on firm growth are illustrative of this point. Researchers have shown that entrepreneurs' desire for rapid or high growth often leads to very different strategic choices from those companies where entrepreneurs do not have this strong need for growth. These differences could significantly influence the scale and scope of the company's operations as well as the types of resources committed to the pursuit of growth in domestic or international markets.

Intangible resources play an important role in the RBV. These resources are widely viewed as a key source of new ventures' competitive advantage at home (Deeds et al., 1999; Itami and Roehl, 1987) or in international markets (McDougall, 1989; Zahra, 2005). For instance, new ventures that have innovative products, control significant patents on breakthrough innovations or hire successful researchers can pursue different strategies (e.g., alliances) from those ventures that do not own these intangibles (Grant, 1996; Itami and Roehl, 1987). New ventures with a strong endowment of intangible resources can also successfully use these resources to get favorable financial deals and obtain the funds they need to develop their own operations. The perceived riskiness of a venture also declines in the minds of creditors and other resource providers once they become aware of the venture's possession of valuable intangible resources such as a new patent on a new technology. This can reduce the "liabilities of newness" that can handicap the development and growth of new ventures.

New ventures can also leverage their intangible resources in ways that augment their gains from their tangible resources, thereby strengthening and protecting its competitive advantage. For example, if the firm develops a proprietary manufacturing process, it can use this process to gain a competitive

advantage by reducing costs or adding features that distinguish its products from those offered by the competition. A new venture that has unique skills in product design can combine these skills with its manufacturing to offer a distinctive new product and compete well in its chosen markets. Of course, the venture can “trade” these skills with others and gain access to distribution channels or other resources that allow it to build a viable market presence.

Assembling and configuring the venture’s tangible and intangible resources requires creativity and foresight (Kazanjian and Rao, 1999). The entrepreneurial qualities of the founders (e.g., alertness, decisiveness and aggressiveness) can serve the venture well at this stage. These qualities help the TMT to see new resource combinations that make the skills of existing competitors irrelevant or make what the new venture has to offer strategically valuable. This is where Schumpeterian competition takes place, shaping the evolution of the venture and its industry. Entrepreneurs learn by doing and from the feedback they get from their markets. This learning could be a valuable resource that allows the venture to recombine its resources differently.

Dramatic changes in tangible and intangible resources could occur following new ventures’ market launch (Bantel, 1998). As these ventures achieve some market acceptance and legitimacy, they might gain access to different types of resources which they can use to change or redefine their businesses as well as where and how to compete. The experience that new ventures gain in assembling and deploying their resources could also shape their future strategic choices. Entrepreneurs might also devise new strategies to organizationally embed their resources and make it difficult for the competition to decipher what they are doing.

## *2.6. Capitalizing on Dynamic Capabilities in Mapping New Venture Strategies*

Recognizing the need for flexibility in assembling and deploying new ventures’ resources (Autio, 2000; Bantel, 1998; Brown and Eisenhardt, 1997), researchers have also studied the development of capabilities within these firms. Capabilities result from the deliberate strategic investments new ventures make in their R&D and other operations. Path dependencies also shape the evolution of these capabilities (Nelson and Winter, 1982; Nelson, 1991; Winter, 2003). These dependencies reflect the fact that prior investments and decisions usually influence and determine the evolutionary trajectories of capabilities, defined as the firm’s integrated set of skills. These skills are gained through experience (as happens in learning by doing), investments in intellectual capital and other strategic activities such as the acquisition of new knowledge from other firms. These skills are further honed, refined and sharpened by use. Entrepreneurs bundle these skills to create distinct capabilities that allow new



ventures to pursue their goals. Some of these capabilities are dynamic, as their texture changes over time and entrepreneurs can use them again and again to pursue a range of strategic objectives. Teece, Pisano and Shuen (1997) define dynamic capabilities as, “The firm’s ability to integrate, build and reconfigure internal and external competences to address rapidly changing environments.”

Some scholars suggest that these dynamic capabilities are at the core of a new venture’s strategic advantages (Eisenhardt and Martin, 2000; Miller, 2003; Teece et al., 1997; Winter, 2003). They give the venture the strategic flexibility necessary to differentiate itself from its rivals. Dynamic capabilities can also be configured quite differently to ensure the strategic variety necessary for competitive distinctiveness. Such novel combinations serve as the foundation of new ventures’ strategic choices. New ventures often have a limited set of capabilities that could be integrated in different ways to ensure a timely response to the demands of the external environment.

The dynamic capabilities perspective has three implications for crafting successful new venture strategies. The first is the need for building and assembling salient capabilities through patient and sustained investments. Entrepreneurs should identify and understand the competitive challenges in their industries and bundle the skills they have to develop their capabilities. New ventures often have a limited set of capabilities and may have to use external sources (e.g., outsourcing and alliances) to offset the limitations of their own skills. Cooperative strategies, discussed elsewhere in this chapter, provide an important way to gain access to these external sources (McGee, Dowling and Megginson, 1995).

A second managerial task for entrepreneurs is to envision novel ways to integrate capabilities. Dynamic capability scholars emphasize integration as a strategic managerial task. For instance, Eisenhardt and Martin (2000) propose, “The firm’s processes that use resources—specifically the processes to integrate, reconfigure, gain and release resources—to match or even create market change. Dynamic capabilities thus are the organizational and strategic routines by which firms achieve new resources configurations as market emerge, collide, split, evolve and die.”

The integration required to build dynamic capabilities demands foresight and creativity, making it difficult for the competition to identify the sources of new ventures’ advantages. Integration challenges entrepreneurs to envision how their industry or market might change and how to deploy their resources accordingly. This integration could be also difficult because the skills that new ventures might have may differ markedly from those imported from external sources. Incoming knowledge and skills often have unique knowledge that is hard to decipher or understand. New ventures might not have the requisite absorptive capacity to identify, acquire, assimilate and exploit the knowledge

embedded in the capabilities they acquire from external sources (Zahra and George, 2002a). For instance, a new venture might participate in an alliance with a well established company to develop a new product. If the venture does not have the internal knowledge or skill to learn from its alliance partner, it misses an opportunity to broaden its knowledge base. Learning is the foundation for new skills and capabilities that define the venture's future strategies (Argote, 1999).

The third managerial task is for entrepreneurs to harvest their organizational learning (Kogut and Zander, 1992; Leonard-Barton, 1995). Entrepreneurs need to develop systems that allow them to capture what their firms have learned, integrate this knowledge into their new venture's strategic planning decisions and infuse it throughout its operations (Helfat and Peteraf, 2003; Leonard-Barton, 1995). Analyzing what the venture has learned from its initial marketing efforts, for example, could reveal important lessons for its production, R&D, logistic and organization design decisions. This learning provides a source of new knowledge for the firm, allowing it to build or revamp new skills. Autio, Sapienza and Almeida (2000) highlight the "learning advantage of newness," positing that new firms often learn quickly because of their organic structures, open and informal communication, lack of strong inertial forces that limit the firm's adaptation processes and the strong involvement of the top leadership in the life of the organization. These advantages of newness, however, hinge on the lead entrepreneur and other TMT members' openness and willingness to formally or informally share their knowledge (Zahra, Ireland and Hitt, 2000). Obviously, new ventures can learn from their successes as well as failures (McGrath, 1995).

The dynamic capabilities perspective places a new venture's knowledge at the center of its activities intended to create strategic distinctiveness, build competitive advantage and gain value. It highlights the importance of having new knowledge and keeping that knowledge fresh through learning by doing, observing the competition, collaboration with other firms and changing the firm's own intellectual base. It also reinforces the important role of the lead entrepreneur and other TMT members in integrating the knowledge generated internally with that acquired from external sources. While internally generated knowledge is frequently difficult for others to understand or imitate and therefore can protect the venture's competitive advantage, entrepreneurs must tap external sources of knowledge and bring that knowledge into their operations. It is the *flow* of knowledge, not simply its stock that makes the difference in a venture's ability to develop and deploy effective strategies (DeCarolis and Deeds, 1999; Deeds et al., 1999). Entrepreneurs need also to tap *different sources* of knowledge, broadening their venture's knowledge base and ensure it explores different strategic options (Miller, 2003).

### 3. COOPERATIVE STRATEGIES

Newly launched ventures do not have to go it alone. Instead, they can collaborate with other ventures and established companies to pursue their goals (McGee, Dowling and Megginson, 1995; Vesper, 1990). Given their limited resources, some ventures have used collaborative (cooperative) strategies to establish their market presence, gain legitimacy, acquire market share and improve performance. Cooperative strategies could substitute for go-it-alone, expensive competitive strategies or could even make these strategies more viable (Dickson and Weaver, 1997). A new venture might use a product development joint venture rather than developing new products on its own. The new venture might use the new product to target new markets using a differentiation strategy and gain a significant market share.

The literature on cooperative strategies is growing rapidly, reflecting their vital importance for new ventures' survival, profitability and growth. This new literature examines the motivation for, structure of, and implications of cooperative strategies. Key themes found in this research are discussed below.

#### *3.1. Antecedents of and Motivation for Cooperative Strategies*

Researchers have discussed the effect of uncertain industry conditions on the use of cooperative strategies. They suggest that high environmental dynamism, unfavorable competitive conditions and complex environments (individually and jointly) encourage new ventures to use cooperative strategies. New ventures may also join forces with others to help develop an industry standard by championing a dominant design, one that is widely accepted in the industry. Of course, the environment does not have to be unfavorable for new ventures to use cooperative strategies. In fact, fast growing industries may favor the use of joint ventures among new ventures to develop and introduce new products that capitalize on this growth.

In addition to challenging environmental conditions, resource limitations might encourage new ventures' use of these strategies in order to gain access to valuable resources or offset the limitations of their own skills. New ventures entering foreign markets often find it useful to cooperate with local or foreign companies already in the market. Cooperation could be multifaceted or limited to one activity along the value chain. New ventures often learn a great deal about the market, distribution, social values and mores through these cooperative ventures. They also learn quite a bit about product design and customization for local markets.

Established corporations also collaborate with new ventures for a variety of reasons. New ventures may control innovative technologies that have the potential to redraw industry boundaries, revise the rules of the competition,

or simply drive existing companies out of business. Collaboration, through joint ventures or alliances, helps established companies to gain valuable information about emerging technologies. Established companies may also gain important insights about emerging organizational forms and business models, which helps them to revise their definitions of their industry's competitive landscape. This learning also opens doors for additional joint new product development and commercialization, a *win-win* situation for new ventures *and* established firms.

Newly launched ventures also learn from established companies' experiences with product commercialization. New ventures typically excel in conceiving and prototyping new product ideas but do not have the organizational or marketing skills to transform these ideas into successful products. By collaborating with established companies, new ventures learn how to organize their operations for this complex but vital task. They can also learn how to assemble the complementary assets required for successful commercialization (Grant, 1998) and obtain the funds needed to continue their operations and survive their initial foray into their markets.

The preceding discussion raises a question: When is it advantageous for new ventures to use cooperative strategies? Of course, it is tempting to say that cooperation is always good and useful. But there are times when it is disadvantageous for new ventures to collaborate. For example, if these ventures are likely to lose their control over their intellectual property and trade secrets to the competition, collaboration is not advantageous. Further, if cooperation reduces the ventures' discretion and degrees of freedom or reduces its strategic flexibility, collaboration may harm these ventures' growth. Likewise, if collaboration results in the loss of new ventures' identity, then it is unwise for entrepreneurs to cooperate with others.

Reflecting on the advantages and disadvantages of cooperative strategies, it becomes clear that the calculus involved is complex. Information about the true intention and characters of potential partners is scarce and often subject to serious causal ambiguity. Relying on partners' track records is helpful but, of course, companies behave in unpredictable ways that depart from their past choices and defy prediction. Therefore, researchers have used game theoretic models to probe different scenarios under which companies should collaborate. Other researchers have used transaction cost theory to explain the conditions favoring the use of cooperative strategies.

Researchers have applied the RBV and dynamic capabilities perspectives in examining cooperative strategies. RBV researchers posit that new ventures could use their tangible and intangible resources to gain access to well establish networks in their industries, build their resource base and create capabilities. Thus, the human, technological and financial resources that new ventures have can serve as a magnet that attracts alliance partners. In turn, alliances enrich these resources and set the foundation for growth. The dynamic

capabilities perspective suggests that the stocks and flows of knowledge (a key and vital resource) gained through alliances provide a foundation for competitive advantage.

The firm's human capital is an important resource that plays a key role in explaining the formation and success of its cooperative strategies. Both the RBV and upper-echelon perspectives emphasize the critical value of the TMT in creating these alliances. The skills, experiences, educational background, values and motivation of new ventures' TMT can profoundly influence access to alliance partners. The social skills that TMT members possess, as well as their personal and professional connections and ties, also influence this process. Personal ties result from social and familial connections and friendships. Professional ties result from job-related connections and relationships. Equally important, the skills and experiences TMT members have in their (as well as other) industries makes it possible to gain access to networks and develop beneficial alliances.

Many new ventures have financial and nonfinancial objectives for their cooperative ventures. For instance, they might use these strategies to build a coalition around an issue of interest to the industry or managers (Lord, 2003) and lobby public policy makers. Some ventures seek to learn about the market and competition, building options for more involved and deeper relationships. Obviously, the importance of the financial and nonfinancial goals and the exact cooperative strategies are likely to differ by company age, size, industry type and senior managers preferences.

### *3.2. Financial Effects of Cooperative Strategies*

Researchers have studied the implications of new ventures' alliances and joint ventures for their successful performance. The studies show that these cooperative strategies could enhance the probability of a venture's survival, improve its perceived market credibility and reputation as it builds its market share. Cooperative strategies also give new ventures access to existing networks in their industries, bringing in new knowledge that fuels innovation and entrepreneurial activities that improve financial performance. Learning from these relationships can also help companies reconfigure its resources differently and devise more efficient ways to make and distribute its products. Over time, these variables could enhance the venture's profitability and growth.

### *3.3. Nonfinancial Effects of Cooperative Strategies*

Alliances and joint ventures are particularly useful in entering new foreign markets and learning about existing technological and innovative paradigms in these markets (Zahra et al., 2000). This is particularly the case when

*social learning* about markets is essential, as happens when new ventures enter markets in countries with very different cultures and norms from their own country. Alliances facilitate social learning, making it easier for entrepreneurs to understand how customers and other key stakeholders make their decisions. This is important because some entrepreneurs have found it difficult to understand the rules that govern consumer decision makers in foreign markets. Alliances with local companies have made this learning possible, helping local entrepreneurs to transfer their tacit knowledge to their Western partners, thus expediting their foreign market. In turn, this has made it possible for Western entrepreneurs to customize their products and better target specific market segments with unique advertising campaigns.

Cooperative strategies also help new ventures to build their absorptive capacity (Zahra and George, 2002a), defined as the capability to spot, identify, acquire, assimilate and use knowledge from external sources (e.g., competition). Enhanced absorptive capacity increases the firm's knowledge base, furthering its ability to harvest the competencies it has while building new capabilities. New ventures that have this capacity learn more easily from their alliance partners, serve their customers, analyze market and supplier feedback and more effectively interpret competitive reactions to their own moves.

Broadening new ventures' absorptive capacity allows them to explore new strategic options. For example, new ventures could combine their internal skills and knowledge gained externally to develop radically new products (Zahra and George, 2000a). These ventures could also use externally acquired knowledge to revise the sequencing, timing and execution of their strategic moves. This facilitates new ventures' improvisation and rapid market responsiveness, helping them to adapt to changes in the market. When absorptive capacity is high, entrepreneurs can also learn by observing changing market dynamics. Even when some of their strategic actions fail, entrepreneurs can quickly regroup and take corrective action. Strategy, thus, becomes a performing art; the more the entrepreneur becomes central to the process and the higher the firm's absorptive capacity, the greater the ability of the entrepreneur to improvise new strategic moves.

One area that is receiving attention in the literature is the effect of cooperative strategies on the evolution of new ventures' dynamic capabilities. By infusing ideas, knowledge and resources into new ventures' operations, the use of cooperative strategies encourages and expedites the development of new capabilities. New ventures' flexible internal systems, proactive management and organic structures make it easier to harness this knowledge, deploy it toward new applications and create growth options.

Several factors are likely to influence the financial and nonfinancial gains new ventures achieve through cooperative strategies. Among the most important are new ventures' experience with these strategies; learning makes a

difference in how new ventures select their partners, manage their relationships with them and learn from the interactions with their partners. The level of trust that might prevail in the relationship between the venture and its partners is also likely to influence these gains. The quality of the partners (e.g., their competence and skills) will also make a major difference. The success of the TMT in creating effective synergies between their firms' competitive and cooperative strategies could also determine firms' gains from cooperative strategies.

#### 4. POLITICAL STRATEGIES

In addition to studying the competitive and cooperative strategies followed by newly launched ventures, researchers have explored their political strategies which focus on the fulfillment of nonprofit making organizational goals. Political strategies seek to legitimize their venture, increase its access to various resources, connect it better to different stakeholders and give it some flexibility in pursuing its cooperative and competitive strategic options. Specifically, researchers have studied the "political" strategies that entrepreneurs might use to influence leading state and governmental agencies and public policy makers, highlighting the importance of proactiveness by entrepreneurs in gaining and maintaining the attention of these groups.

Hillman and Hitt (1999) note that attention in the literature to the political strategies companies follow is recent, even though their importance has long been recognized (Keim, 1981). Public policy issues have not been well integrated into business school curricula and they have been traditionally left to political scientists and other interested parties. Now that countries, states and agencies have come to appreciate the vital importance of entrepreneurship, greater effort is being devoted to understanding the various ways new ventures can capitalize on this momentum. In the U.S., there is considerable interest in fostering the creation of new firms as a means of revitalizing the economy and strengthening its global standing. Job creation is an important issue that cuts across different political party lines. Almost every state has created multiple programs to help aid the development and growth of new firms, especially those in high technology industries. Public discussion has focused on finding ways to help these companies gain access to funds, incorporate modern technology, improve their human capital and internationalize their operations.

Outside the U.S., interest in entrepreneurship and new ventures has been great. In one country after another, there has been a surge of interest in creating the infrastructure that supports the birth, incubation and growth of new firms. This interest has compelled new business owners to work with others to sustain supportive public policies that encourage and reward entrepreneurship. New business founders have collaborated also to map out strategies that aim to change trade and tax policies that stifle individual initiatives.

Widenbaum (1980) suggests that companies and industries vary in their approaches to public policy. Some passively react to these policies where others proactively participate in the political arena, hoping to shape evolving decisions. New ventures are learning to be more proactive in the policy arena—creating and following different strategies intended to create a more favorable business climate for their operations. These firms use a wide range of tactics that include membership in trade associations and contributions to lobby groups that promote the interests of new ventures.

Hillman and Hitt (1999) suggest that the political strategies firms use fall into three categories: information, financial incentive and constituency-building. *Information* strategies center on giving public policy makers the information they need to develop their plans for the industry's future. Lobbying, special studies and reports, and provision of expert witnesses are examples of the tactics that companies use in this process. *Financial incentive* strategies target political decision makers by giving them financial resources through contributions to candidates or political parties, and providing volunteers to work on special issues. These contributions aim to win the support of public policy decision makers. *Constituency building* strategies seek to create awareness among interested groups and galvanize their support for given causes (Lord, 2003). Companies use several tactics in this process. These include press conferences, public relations, advertising campaigns and town meetings.

Short on resources and pressured to work hard to survive, new ventures clearly stand to gain a great deal from using political strategies. Yet, owners of these companies have very diverse agenda, working styles, time horizons and incentives. These factors can limit these companies' ability to proactively engage public policy makers. A firm's limited resources, limited staff and owners' busy schedules are additional barriers that can limit the use of these strategies. Hillman and Hitt (1999) suggest, however, that the more the firm is dependent on the government (and general environment) for survival and the broader its market definition, the more likely it will make use of political strategies. Indeed, the more the serious the issue at stake, the more likely that these firms will follow political strategies (Keim, 1981).

Recent research on social entrepreneurship also offers important clues about the importance of "political" strategies for new ventures' successful performance. *Social entrepreneurship* means different things to different people. This label, however, highlights the intersection of entrepreneurial risk taking and a commitment to social responsibility, suggesting that some entrepreneurial ventures are created to do some good rather than simply making money and enriching their owners (Alvord et al., 2004; Harding, 2004). Social entrepreneurs have created companies to address specific needs in their own communities, investing in environmentally friendly technologies and collaborating with nonprofit organizations in targeting issues of interest to society and its citizens.



Academic programs related to social entrepreneurship have grown rapidly in the U.S., responding to the need for a greater understanding of the challenges facing these newly created firms. For instance, we know little about the factors that motivate entrepreneurs to develop these companies, how these entrepreneurs choose the social issues or causes they emphasize and how they integrate the profit-making and the “social responsiveness” goals in their strategies.

To be sure, social entrepreneurs are interested in making profits and creating wealth (Harding, 2004). Yet, they have a different mission—one that values social responsiveness. Achieving an effective balance between the demands of responsiveness and making a profit requires creativity and skills. Research is needed to understand how new firms achieve and maintain this balance and why certain ventures might be more adept than others at doing this. To be sure, some social entrepreneurs are believers in the causes they champion and therefore create their companies to ensure attention to these causes. These entrepreneurs would not see their commitment to social issues as a part of a deliberate political strategy. Attention to social issues can enhance a new venture’s legitimacy, increase its market presence and ensure its sustainability.

The political strategies new ventures will use depend greatly on the issues at hand. The more complex the issue and more threatening it is to the firm’s survival, the more likely that the TMT will become more politically proactive. Similarly, when the issue has long-term consequences for the firm and its operations (e.g., internationalization), the more likely managers will be to use proactive political strategies. Of course, companies in different environments or different resource bases are likely to differ also in the tactics they follow in effecting their political strategies.

## 5. KEY CONTINGENCIES IN SELECTING NEW VENTURE STRATEGY

There is no perfect strategy that will ensure the success of every firm in every industry. Rather, each new venture needs a strategy that best responds to the conflicting demands of its external environment, ensures the achievement of the goals of its lead entrepreneur and TMT, capitalizes on the skills and composition of its TMT, and exploits the uniqueness of the resources and opportunities associated with its own heritage and origin. This section addresses these issues in turn.

### *5.1. The External Environment and New Venture Strategy*

Early research has examined the effect of the industry structure and external environment on the strategic choices new ventures make (Chandler and Hanks, 1994; Child, 1972; Covin et al., 1990; Dean and Meyer, 1996;

McDougall, Robinson and DeNisi, 1992; Robinson, 1999; Robinson and McDougall, 1998; Sandberg, 1986; Shepherd and Shanley, 1998). Researchers have also examined the effect of the *perceived* characteristics of the external environment on new venture strategy, arguing that how entrepreneurs view and interpret their external environment influences the strategic choices they make for their new ventures regarding the scope and scale of their operations and timing and sequencing of different strategic moves. Researchers emphasized three dimensions of the external environment as major sources of influence on new venture strategy: dynamism, hostility and heterogeneity.

*5.1.1. Environmental Dynamism* Dynamism refers to the extent of changes that are taking place in the external environment, the magnitude of these changes and their direction (Ginsberg and Venkatraman, 1985). These changes make planning difficult, if not futile. Instead, these changes encourage flexibility and innovativeness (Zahra, 1996b; Zahra and Bogner, 2000). In dynamic environments, entrepreneurs have to improvise (Moorman and Miner, 1998) as they seek to adapt to an ever-changing competitive landscape. In this environment, entrepreneurs often encounter new issues that demand their immediate attention. New ventures also need to pursue product and process innovations and build their market positions as well as gain a competitive advantage. Further, these ventures have to upgrade their products quickly in order to keep up with shifts in technological and market conditions. In turn, this requires strong R&D, engineering and manufacturing capabilities. When entrepreneurs view their environment as dynamic, their R&D activities emphasize transforming technological discoveries into marketable products and goods that new ventures could commercialize quickly and efficiently. Product innovation, radical or incremental, is conducive to higher profitability and growth because dynamism creates opportunities that entrepreneurs and their firms exploit. Cooperative strategies make this possible (Dickson and Weaver, 1997).

There are upper limits on what new ventures can achieve through innovation in dynamic environments. Pressured to continuously innovate, any advantages these ventures might have are at best short lived (Bahrami and Evans, 1989; Brown and Eisenhardt, 1997). With successive innovations, any financial gains are likely to evaporate as the firm has to reinvest in innovation simply to stay alive. New ventures need also to be active in marketing and distributing their products and building their name recognition. These are time consuming and costly activities that require considerable resources, which are often in short supply among newly launched ventures. As already noted, this should encourage new ventures to seek collaborative relationships such as alliances and joint ventures with other companies to develop and market new products. Research reveals that new ventures are likely to enter into collabora-

tive relationships when they see their environments as dynamic (McGee et al., 1995).

*5.1.2. Environmental Hostility* Hostility indicates the perceived unfavorability of a firm's external environment in terms of the sources that exist or the support it can provide the organization. This unfavorability arises from the existence of strong and well endowed competitors who have strong capabilities that they could use to preempt new ventures' entry (Ginsberg and Venkatraman, 1985). Hostility emanates also from the structure of the venture itself such as a heavy concentration in industry sales, maturity of the technology, the capital intensity of operations or the need for high start-up costs. These variables make the successful launch of new ventures difficult and increase the riskiness of their operations. Declining demand or maturing technologies may not support the survival or financial performance of these ventures.

Hostility also arises from technological changes that alter industry structures and bring about fundamental shifts in manufacturing and marketing regimes in the industry (Ginsberg and Venkatraman, 1985). Some of these changes are favorable, creating opportunities that newly launched ventures could successfully exploit. Yet, other changes can damage the skill bases of existing companies or alter the very fabric of the industry, encouraging new ventures to develop ways to better reposition themselves. Pressures mount as new ventures prove the viability of their technologies and products, prompting the entry of new and established companies from adjacent industries. On the other hand, market attractiveness is important for successful new venture performance (Chandler and Hanks, 1994; Dean and Meyer, 1996; Ginsberg and Venkatraman, 1985; Robinson, 1999).

Researchers also have examined the implications of the perceived environmental hostility on new ventures' competitive strategies. Research shows that the perceived hostility of the environment compels entrepreneurs to carefully choose a viable market niche which they can develop and grow their business. These niches are market spaces that have been overlooked by existing companies or other new ventures. New niches might also come into existence as different technologies or as segments converge (Anderson and Tushman, 1990). For instance, in the hay days of the dot-com era, many new ventures were able to capitalize on this convergence by offering integrated "one-stop shopping" business models, thus serving as consolidators of different services and product offerings. Along with these new companies came boutique firms that offered a wide assortment of specialized services and goods, filling vacant niches in the industry.

Perceived environmental hostility can also encourage new ventures to specialize and develop a set of unique capabilities that give them a competitive advantage. Yet, as hostility increases, entrepreneurs might decide to flee the

industry and explore other options. One of the major mistakes that some entrepreneurs make is the lack of strategic commitment (i.e., sustained investment in building unique capabilities) and absence of strategic coherence (i.e., pursuing too many unrelated activities that do not reinforce each other). While the need for flexibility is paramount in this environment, decisiveness about areas to pursue and levels of investments are key pathways to organizational survival and financial success. Entrepreneurs need also to use collaborative strategies (e.g., alliances) and shield their new ventures from the adversity of their hostile external environments.

*5.1.3. Environmental Heterogeneity* The perceived heterogeneity of the environment can also influence new venture strategy. Heterogeneity refers to the extent to which entrepreneurs view their industries and markets as encompassing multiple and distinct segments with different needs and requirements for success (Ginsberg and Venkatraman, 1985). A heterogeneous environment is complex because its dimensions are often interconnected and a change in one dimension often evokes serious changes in other dimensions, sometimes in unforeseen and unfavorable ways. This complexity challenges entrepreneurs' cognitive abilities to select those parts of the industry in which they want to participate (their "served markets"). Given that some entrepreneurs are prone to oversimplification (Simon and Houghton, 2002), this can lead to serious miscalculations and errors in defining the served market and the competition.

New ventures, being limited in resources and skills, may follow a sequential market entry strategy when the industry is heterogeneous. They initially target a given market or niche and once new ventures are well positioned in that market they proceed to build their presence in a new market arena. Thus, some ventures that are launched using a Defender strategy might venture into a new segment, adopting an "Analyzer" or "Prospector" strategy (Miles and Snow, 1978). This approach helps new ventures to ration their resources, overcome the liability of newness and establish enduring and viable contacts that they can later use to expand their operations.

In a heterogeneous environment, entrepreneurs should broaden their search for opportunities without losing track of their strategic thrust. Searching for opportunities could be done informally through proactive networking. The information gleaned from these contacts provides important clues about additional market opportunities that new ventures can target. The search for opportunities could also be done more formally through market research, R&D activities, or competitive analysis. Still, entrepreneurs should achieve some degree of strategic coherence in the strategies they follow, otherwise risk diluting their resources. This means they should have a clearly articulated and well integrated set of strategic choices, as happens when the firm uses

the Analyzer or Prospector strategies. One way to do so is to seek business opportunities that lie at the intersection of two or more market segments.

The use of political strategies is also likely to increase when the venture's external environment is viewed as uncertain (e.g., dynamic, hostile or heterogeneous). This uncertainty is likely to encourage entrepreneurs to join trade associations as they lobby legislators, work with regulatory bodies considering new policies that govern the industry, or other firms attempting to position the industry and its products more favorably in the minds of the general public. These time-consuming activities, however, can lower the firm's uncertainty, giving it a better perspective on the future trajectory of the industry's evolution. Over time, these political activities can lower the firm's operating costs and stabilize the supply of raw material and other resources, thus increasing the odds of the new venture's survival and successful performance.

### *5.2. Founding Entrepreneurs' Goals and Intentions*

The payoff from new venture strategies depends on the goals of the founder and the TMT. These goals reflect the entrepreneurs' family and business histories, needs and values, and psychological make-up such as the need for recognition (Moran, 1998). These goals provide the motivation to launch new ventures in the first place (Birley and Westhead, 1994). Some entrepreneurs create their companies out of necessity, hoping to create employment opportunities for themselves and their families. These entrepreneurs might feel blocked out of the employment opportunities in existing companies because of their skills, race, age or other variables. The Global Entrepreneurship Monitor (GEM) research program reveals that necessity is a major reason for creating new firms across the globe, with certain developing countries leading others in this regard (Reynolds et al., 2001). Necessity-based new ventures do not grow as fast, probably because owners do not have the incentives, skills and resources needed to grow. They may also want to retain control over their companies' operations.

Other entrepreneurs create companies to capitalize on their skills, resources and opportunities. Such "opportunity-based" entrepreneurs are more prevalent in advanced and growing economies where the infrastructure is well developed, access to resources is great, supporting services are common and opportunities are plentiful. In this context, venture growth creates wealth for owners and their families. However, whether the firm pursues growth or not depends on the owners' own goals. Research indicates that owners' intentions to grow the firm is an important variable in influencing its strategic choices such as the geographic scope of the business, the breadth of its business definition, aggressiveness of its marketing and distribution, and the use of

strategic alliances (Bamford, Dean and McDougall, 2000). The “intention to grow” shapes entrepreneurs’ strategic choices and how they execute their plans as well as the way they assemble the TMT, obtain and allocate different resources and sequence market entry (Wiklund and Shepherd, 2003). They also determine whether new ventures make use of the various cooperative and political strategies discussed in this chapter.

Growth intentions also determine the level of energy that entrepreneurs exert in pursuing the strategic options that might be available for their new ventures. Entrepreneurs who are eager to grow their operations rapidly might devote more time to connecting with others, seeking new business opportunities at home or abroad, exploring alliance options and opening new international markets (Davidsson and Wiklund, 2000; Eisenhardt and Schoonhoven, 1990; Wiklund and Shepherd, 2003). These entrepreneurs might devote more energy to day-to-day management issues or work harder at motivating their employees to be productive. These activities can spell the difference in the gains that a firm achieves from its strategic choices.

### *5.3. TMT Composition and Skills*

Most new ventures are established and led by a team (Birley and Stockley, 2000). Research indicates that the composition and skills of the newly launched ventures’ TMT can influence its strategic choices (Eisenhardt and Schoonhoven, 1990; Ensley, Pearson and Amason, 2002). The TMT consists of a firm’s most senior managers who shape its market definition, resource allocation and choices of competitive weapons. The values and aspirations of the TMT often determine the scope and scale of a new venture’s operations (Hambrick and Mason, 1984). A TMT that is dominated by high achievement-oriented members often values growth, rapid expansion and aggressive market entry in domestic or international markets. When the TMT members differ in their educational backgrounds or business experiences, they emphasize variety, innovativeness and risk-taking strategies (Bantel and Jackson, 1989; Hambrick and Mason, 1984). Ventures that are led by such teams are also more likely to explore a wide range of strategic choices, showing greater flexibility than their competitors in shifting from one strategy to the next. The same could be said about the variety of the educational backgrounds of the TMT; the greater that variety, the more likely the new venture will stress innovativeness and produce differentiation in their operations and strategies. Different educational backgrounds often lead team members to see and interpret their environments quite differently and develop different strategies that capitalize on unique opportunities in their environments (Hambrick and Mason, 1984). Different educational backgrounds and functional experiences also lead different TMT

members to spot different types of opportunities in their industry, domestically and internationally.

Finally, differences in TMT members' ages can influence new venture strategies in significant ways (Birley and Stockley, 2000; Hambrick and Mason, 1984). Teams dominated by younger members are more likely to value experimentation and innovation and encourage risk taking, spurring domestic and international growth. These members are also likely to be ambitious and pursue aggressive strategies. In fact, younger TMTs usually lead Prospector organizations where innovation and expansion are the two dominant strategic thrusts. Conversely, older managers control Defender organizations where stability and incremental innovation are valued (Zahra and Pearce, 1990).

The diversity of the top management team can help to connect the new ventures to different power centers in the industry, facilitating the acquisition of important information about potential industry trends. These connections make it easier to gain access to resources and do so reliably and relatively inexpensively. In turn, this could reduce the firm's costs. Connections also make it possible to enter into alliances. These beneficial connections can also help the firm join networks in their industry and effectively execute their political strategies.

#### *5.4. Origin and New Venture Strategy*

Research has also documented the implications of new ventures' origin for their strategic choices. New venture origin indicates whether the firm is independently created and managed or it has been formed within a well-established organization (MacMillan and Day, 1987; Zahra, 1996a). Well-established organizations have experiences, networks and resources that can be leveraged to support the development and subsequent growth of their ventures (Miller Camp, 1985). Still, these ventures might be constrained by the parent's existing systems that limit managers' discretion and curtail their innovativeness. Even when corporate-sponsored ventures are given autonomy, their decisions are subject to the approval and support of the parent corporation. This can slow down these ventures' decision-making processes. Independently owned ventures often experience severe resource constraints, poor name recognition, lack of connections, limited access to viable network relationships and other liabilities that arise from their newness. Given these potential differences, some researchers have suggested that independent and corporate-sponsored new ventures would select different market niches, choose different strategic weapons and seek different sources of competitive advantage. These differences could lead to variability in new venture performance (Zahra, 1996a).

Some studies have shown that corporate-sponsored ventures tend to be broader in their market scope, spend more on advertising and marketing,

have easier and greater access to distribution of market channels and are larger in their organizational size (measured by full-time employees) than independent new ventures (Zahra, 1996a). Research has been less conclusive on the differences between independent and corporate ventures in their spending on R&D and innovation activities. However, some studies tentatively suggest that independent new ventures usually emphasize creating and introducing radically new products in order to differentiate themselves from the competition. Corporate ventures tend to be more prolific in their product introductions and upgrading them than independent ventures (Zahra, 1996a).

Research has been less conclusive on the differences between independent and corporate ventures in their financial performance (for reviews, see Biggadike, 1979; Shrader and Simon, 1997; Zahra, 1996a). Some studies have theorized that, because of the rich resources of their sponsors, corporate-sponsored ventures would outperform those created by individual entrepreneurs. This proposition has not been well supported in the literature, as some studies have found no discernible differences between the two types of ventures. Others have discovered that independent ventures outperformed those created by well-established corporations (for a review, see Shrader and Simon, 1997). These findings reflect the possibility that some established companies are slow to create and launch their ventures, giving independent entrepreneurs time and space to develop their firms and position them well in their markets. Established companies tend to be bureaucratic and tightly control their ventures, limiting managers' discretion and ability to respond in a timely fashion to changing industry forces. Independent ventures are more innovative and entrepreneurial in their strategies than their corporate counterparts.

The above observations should be interpreted with caution because prior studies on the effects of new venture origin suffer from survivor bias. Most of the studies have also been cross-sectional, making it difficult to consider alternative explanations of the findings reported to date. One consequence of these shortcomings is that they do not know how and where independent and corporate ventures compete or the extent of the differences in their strategies. We also cannot tell from the literature if independent new ventures have higher or lower mortality rates. And, if there are significant differences, we do not know much about the factors that explain variability in these survival rates over time. These are important issues that deserve analysis in future studies, especially as more corporations launch new ventures to capitalize on the changes in their industries.

An issue that deserves additional research is the nature of the relationships that might develop between independent and corporate-sponsored new ventures. Independent entrepreneurs frequently retain formal and informal links with established companies (Bhide, 2000), having gained considerable work experience and having learned a great deal about the industry and the compe-



tion. This experience often shapes the strategic choices these entrepreneurs make when they launch their companies, including any formal and informal linkages they develop with their former employers. These relationships, in turn, provide important means of learning and gaining access to information about the competition, the market and other vital resources that could influence the scope and scale of the strategic choices of the entrepreneur.

Researchers have frequently treated independent and corporate ventures as rivals who seek each others' destruction. While these ventures compete with one another, they also collaborate. For instance, corporate venture capital (CVC) has become a vital source of funding for independent ventures to develop and test their new products, giving established companies (investors) important clues about the evolutionary trajectory of new technology and how it may shape competitive dynamics in their own markets. This information could be valuable for established companies as they start their own in-house ventures. In addition, as established corporations become cognizant of the new technologies and business models that independent ventures have developed, they may form beneficial alliances to jointly develop and introduce their products. These relationships might also set the stage for established companies to acquire these ventures and integrate them into their ongoing operations, renewing their operations by creating new avenues for growth.

## 6. THE EFFECT OF THE ENVIRONMENT ON THE STRATEGY-PERFORMANCE RELATIONSHIP

Some researchers have noted that the financial effects of a venture's competitive strategy depend on the characteristics of the external environment (Buzzell and Gale, 1987; Dickson and Weaver, 1997; Miles and Snow, 1978). This means that the environment could moderate the relationship between the competitive strategy and a firm's performance. Research using the profit impact of market strategy (PIMS) has supported this proposition (Buzzell and Gale, 1987), revealing that some strategic choices positively influence the firm's profitability and market share in a given environment but have the exact opposite effect or even no effect in other environments. Despite the plausibility of these findings, they were derived from data based on strategic business units of well-established companies for whom resources were not a major concern. Many of these units also competed in maturing or declining industries, possibly limiting variability in their competitive strategies.

Other research, however, supports the moderating effect of the external environment on the relationship between competitive strategy and performance (McDougall et al., 1992; Zahra, 1996b; Zahra and Bogner, 2000). For example, Miles and Snow's (1978) Defender strategy is likely to lead to higher profitability in stable rather than dynamic environments. The opposite is true of

Prospector strategies. Porter's (1980) low-cost strategy is conducive to higher profits in stable environments but a product differentiation strategy generates higher profits when the business environment is dynamic. Zahra and Kirchoff (2005) found that some of strategic choices new ventures make about their technology have a positive effect in high technology industries but have a negative or no effect in low-technology industries. These results support the view that each industry has its requirements for successful performance (Grant, 1996, 1998; Porter, 1980) and entrepreneurs should understand these variables and craft their new ventures' strategies accordingly.

Reaping the financial gains associated with new venture strategies requires entrepreneurs to analyze the industries in which they plan to launch their ventures. Unfortunately, entrepreneurs do not always undertake systematic or comprehensive analyses of their industries, possibly blinding them to pending threats and challenges. Some entrepreneurs rely more heavily on talking to their major customers or suppliers or friends. While these groups might provide important insights about the industry and its potential evolution, entrepreneurs may not have full access to diverse data that challenge their assumptions. Also, even when they conduct industry analyses, some entrepreneurs do this on an *ad hoc* basis, raising the odds of faulty conclusions or serious strategic errors. Entrepreneurs are prone to oversimplify complex facts, overemphasize a few factors in their environmental and industry analysis, dismiss contradictory information with which they are not comfortable and generalize based on limited numbers or cases (Simon and Houghton, 2002). These variables could render the results of industry analyses meaningless. These problems are compounded by the fact that new ventures do not have the resources, skills, experiences or staff to conduct their independent analyses. Further, with founders firmly in command of their ventures, few in the organization are able or willing to challenge their interpretation of the industry or the forces that drive its evolution.

## 7. FUTURE RESEARCH AGENDA

Throughout this chapter, I have highlighted several areas worthy of future exploration and research. In this section, I will pause to reflect on several additional issues that deserve recognition and investigation.

### 7.1. *Competitive Strategies*

This review of the literature suggests several research questions that require attention in future studies, especially the role of resources and dynamic capabilities.

*7.1.1. The Role of Resources* As noted earlier in the chapter, a vast body of research now exists on the RBV and its usefulness in the context of new ventures. Results from this research tentatively suggest that new ventures that embed their various resources in their products, systems, processes or operations can create a sustainable source of competitive advantage. Presently, we do not know much about the activities that entrepreneurs undertake to embed their resources into their organization or the consequences of the various approaches that entrepreneurs might use for the content of their new venture strategies. Changes in the membership of the TMT or the pattern of interaction among them may also alter the way they assemble, configure and deploy their various resources. Here, too, little research currently exists on how these variables are likely to influence potential changes in the strategies new ventures might use.

Similarly, little empirical research has documented the differences in tangible and intangible resources between new ventures and established companies. This research has also failed to articulate the key differences between corporate and independent ventures in how they combine their different tangible and intangible resources. Given that these firms often battle each other for survival, researchers should document the differences that might exist between them in their tangible and intangible resources. They also need to link these differences to new ventures' success or failure and growth. The characteristics of new ventures' industry might moderate the relationships between resources and new venture strategy and between that strategy and venture performance, as discussed earlier.

*7.1.2. Dynamic Capabilities* As noted earlier, there is considerable interest today in the strategy and entrepreneurship literature in understanding the nature and effects of dynamic capabilities. Despite this interest, several issues await future research. For example, how different are new ventures' dynamic capabilities from those observed in well-established companies? Given that dynamic capabilities result from multiple resources and skills and managerial processes, it would be logical to assume that significant differences exist between new ventures and established companies in these capabilities and perhaps the variety of capabilities they have. The variables that trigger changes in these capabilities also deserve attention, given the differences that might exist between established and new ventures' internal and managerial decision-making processes. Other questions to explore in future research are: How do new ventures keep their capabilities dynamic? Why do certain capabilities lose their strategic relevance? What do new ventures do with the assets dedicated to dated capabilities?

### 7.2. *Cooperative Strategies*

As noted, research on cooperative strategies has increased in recent years, focusing on the implications of these strategies for new venture performance. Even though many researchers have touted the importance of cooperative strategies, some have explored their negative and dysfunctional effects. Cooperative strategies can leak important information to rivals about a new venture's operations, diluting its competitive advantage. The need for coordination can also raise costs and slow down entrepreneurs' decisions. There is always the risk of partners' opportunism (Dyer and Singh, 1998), where a well-resourced corporation could learn enough from a venture to preempt its new product introductions. Entrepreneurs often have to safeguard against this opportunism. Even when trust develops in a cooperative alliance, opportunism cannot be totally eliminated. Finally, alliances may become a prelude to takeover, even hostile, bids.

There is a voluminous body of research on cooperative strategies, the bulk of which is based on the experiences of larger companies. Research into the role of cooperative strategies in new ventures' success is also growing. Still, research is needed to understand how and when alliances and other cooperative strategies help new ventures to learn and assemble the capabilities they need to survive and achieve profitability. Further, to date, researchers have limited their attention to the short-term consequences of cooperative strategies for new ventures' performance. It is also unclear from the literature as to how these strategies influence ventures' transitions from one stage of their life cycle to the next.

One area that requires further attention is the alliances formed between corporate and independent new ventures. Given that these ventures compete in the same market, the dynamics and structure of these alliances might be different from other types of alliances. Further, independent ventures are often created earlier than their corporate counterparts and might have important knowledge and skills corporate-sponsored ventures do not possess. The implications of this knowledge asymmetry for the types of alliances formed, durability of these alliances and sharing of knowledge among partners deserve future study.

### 7.3. *Political Strategies*

Findings from research on social entrepreneurship and new ventures' social responsibility suggest some exciting questions to be explored in future studies. What is the relative importance of competitive vs. cooperative vs. noncompetitive strategies in determining new venture performance? Does this importance vary across venture and industry types? How do entrepreneurs balance these three facets of strategy as they seek to (re)position their newly

launched ventures? What role does “noncompetitive” strategy play in making competitive and cooperative strategies more viable? Does this role vary from one industry setting to the next?

#### *7.4. The Three Dimensions of New Venture Strategy*

The relationships between the competitive, cooperative and political facets of a firm’s new venture have not been closely examined. We need to know more about the content of new ventures’ political and cooperative strategies. Researchers from fields as diverse as strategy, economics, finance and international business have studied cooperative strategies, offering multiple theories for their formation and effects. These theories could be useful in studying new venture strategies. We know far less about the political strategies that new ventures use and it would be useful to document and classify them.

How and when entrepreneurs use political and cooperative strategies remains unclear in the literature. Do entrepreneurs use cooperative and political strategies in lieu of (or complement to) competitive strategy? How do these strategies influence short- and long-term new venture performance? Also, empirically, when do these strategies complement and substitute each other? Are certain entrepreneurial qualities related to the use of these three strategies? Do new ventures use these strategies differently at different stages of their life cycle? Do corporate ventures vary from their independent counterparts in their focus on these strategies? If so, what are the consequences for organizational performance?

## 8. CONCLUSION

Newly launched ventures face an uncertain future. Careful planning and organizing in the pre-launch stage can reduce this uncertainty. However, as these ventures are launched, they have to manage the powerful and unpredictable forces of the market and competition. They have to learn to maneuver and fly their way into existence and survival. They need also to be clever in selecting their markets, assembling their resources, defining their strategic thrusts and determining their competitive weapons. These ventures need to learn and conceive of new ways to compete. By crafting creative competitive, cooperative and political strategies, entrepreneurs give their ventures the power to survive, make a profit and grow.

## NOTE

<sup>1</sup> I acknowledge with appreciation the valuable comments and guidance of Simon Parker, two anonymous reviewers and Patricia H. Zahra.

## REFERENCES

- Aldrich, H. (2000). *Organizations Evolving*. Sage, New York.
- Alvord, S., Brown, L. D. and Letts, C. (2004). Social Entrepreneurship and Societal Transformation. *Journal of Applied Behavioral Science* 40: 260–283.
- Anderson, P. and Tushman, M. (1990). Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change. *Administrative Science Quarterly* 35: 604–633.
- Argote, L. (1999). *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic Publishers, Norwell, MA.
- Andersson, S. (2004). Internationalization in Different Industrial Contexts. *Journal of Business Venturing* 19: 851–875.
- Autio, E. (2000). Growth of Technology-Based New Firms. In *The Blackwell Handbook of Entrepreneurship*, D. L. Sexton and H. Landstrom (Eds), pp. 329–347. Blackwell Publishers Ltd., Oxford, UK.
- Autio, E., Sapienza, H. J. and Almeida, J. G. (2000). Effects of Age at Entry, Knowledge Intensity, and Imitability on International Growth. *Academy of Management Journal* 43 (5): 909–924.
- Bahrami, H. and Evans, S. (1989). Strategy Making in High-Technology Firms: The Empiricist Mode. *California Management Review* 31: 107–128.
- Bain J. S. (1959). *Industrial Organization*. John Wiley and Sons, New York.
- Bamford, C. E., Dean, T. J. and McDougall, P. P. (2000). An Examination of the Impact of Initial Founding Conditions and Decisions upon the Performance of New Bank Start-ups. *Journal of Business Venturing* 15: 253–277.
- Bantel, K. A. (1998). Technology-Based, “Adolescent” Firm Configuration: Identification, Context, and Performance. *Journal of Business Venturing* 13: 205–230.
- Bantel, K. A. and Jackson, S. E. (1989). Top Management and Innovations in Banking: Does the Composition of the Top Team Make a Difference? *Strategic Management Journal* 10: 107–124.
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management* 17 (1): 99–120.
- Barney, J. (2001a). Resource-Based Theories of Competitive Advantage: A Ten Year Retrospective on the Resource-Based View. *Journal of Management* 27 (6): 643–650.
- Barney, J. B. (2001b). Is the Resource-Based “View” a Useful Perspective for Strategic Management Research? Yes. *Academy of Management Review* 26 (1): 41–56.
- Barney, J. B., Wright, M. and Ketchen, D. (2001). The RBV of the Firm: Ten Years after 1991. *Journal of Management* 27: 625–641.
- Bhide, A. (2000). *The Origin and Evolution of New Business*. Oxford University Press, New York.
- Biggadike, E. R. (1979). *Corporate Diversification: Entry, Strategy, and Performance*. Harvard University, Boston, MA.
- Birley, S. and Stockley, S. (2000). Entrepreneurial Teams and Venture Growth. In *The Blackwell Handbook of Entrepreneurship*, D. L. Sexton and H. Landström (Eds), pp. 287–307. Blackwell, Oxford, UK.
- Birley, S. and Westhead, P. (1994). A Taxonomy of Business Start-up Reasons and their Impact on Firm Growth and Size. *Journal of Business Venturing* 9: 7–31.
- Brown, S. L. and Eisenhardt, K. M. (1997). The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations. *Administrative Science Quarterly* 42 (1): 1–34.
- Buzzell, R. D. and Gale, B. T. (1987). *The PIMS Principles—Linking Strategy of Performance*. The Free Press, New York.
- Carter, N. M., Gartner, W. B. and Reynolds, P. D. (1996). Exploring Start-up Event Sequences. *Journal of Business Venturing* 11: 151–166.

- Carter, N. M., Stearns, T. M., Reynolds, P. D. and Miller, B. A. (1994). New Venture Strategies: Theory Development with an Empirical Base. *Strategic Management Journal* 15: 21–41.
- Chandler, G. N. and Hanks, S. H. (1994). Market Attractiveness, Resource-Based Capabilities, Venture Strategies, and Venture Performance. *Journal of Business Venturing* 9: 331–349.
- Child, J. (1972). Organizational Structure, Environment and Performance: The Role of Strategic Choice. *Sociology* 6: 1–22.
- Covin, J. G., Slevin, D. P. and Covin, T. J. (1990). Content and Performance of Growth-Seeking Strategies: A Comparison of Small Firms in High- and Low-Technology Industries. *Journal of Business Venturing* 5: 391–412.
- Davidsson, P. and Wiklund, J. (2000). Conceptual and Empirical Challenges in the Study of Firm Growth. In *The Blackwell Handbook of Entrepreneurship*, D. L. Sexton and H. Landstrom (Eds), pp. 26–44. Blackwell Publishers Ltd., Oxford, UK.
- Dean, T. J. and Meyer, G. D. (1996). Industry Environments and New Venture Formations in U.S. Manufacturing: A Conceptual and Empirical Analysis of Demand Determinants. *Journal of Business Venturing* 11: 107–132.
- DeCarolis, D. M. and Deeds, D. L. (1999). The Impact of Stocks and Flows of Organizational Knowledge on Firm Performance: An Empirical Investigation of Biotechnology Industry. *Strategic Management Journal* 20: 953–968.
- Deeds, D. L., DeCarolis, D. and Coombs, J. (1999). Dynamic Capabilities and New Product Development in High Technology Ventures: An Empirical Analysis of New Biotechnology Firms. *Journal of Business Venturing* 15: 211–229.
- Dickson, P. H. and Weaver, K. M. (1997). Environmental Determinants and Individual-Level Moderators of Alliance Use. *Academy of Management Journal* 40: 404–425.
- Dyer, J. H. and Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Academy of Management Review* 23 (4): 660–679.
- Eisenhardt, K. M. and Martin, J. A. (2000). Dynamic Capabilities: What are They? *Strategic Management Journal* 21 (10–11): 1105–1121.
- Eisenhardt, K. M. and Schoonhoven, C. B. (1990). Organizational Growth: Linking Founding Team, Strategy, Environment, and Growth among U.S. Semiconductor Ventures, 1978–1988. *Administrative Science Quarterly* 35 (3): 504–529.
- Ensley, M. D., Pearson, A. W. and Amason, A. C. (2002). Understanding the Dynamics of New Venture Top Management Teams: Cohesion, Conflict, and New Venture Performance. *Journal of Business Venturing* 17: 365–386.
- Fleisher, C. S. and Bensoussan, B. E. (2003). *Strategic and Competitive Analysis: Methods and Techniques for Analyzing Business Competition*. Prentice Hall, Upper Saddle River, NJ.
- Ginsberg, A. and Venkatraman, N. (1985). Contingency Perspectives of Organizational Strategy: A Critical Review of the Empirical Research. *Academy of Management Review* 10: 421–434.
- Grant, R. M. (1996). Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal* 17: 109–122.
- Grant, R. M. (1998). *Contemporary Strategy Analysis*, 3rd edn. Blackwell Publishers Inc, Massachusetts, US.
- Hambrick, D. C. and Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of its Top Managers. *Academy of Management Review* 9 (2): 193–206.
- Harding, R. (2004). Social Enterprise: The New Economic Engine? *Business Strategy Review* 15 (4): 39–44.
- Helfat, C. E. and Peteraf, M. A. (2003). The Dynamic Resource-Based View: Capability Lifecycles. *Strategic Management Journal* 24 (10): 997–1010.
- Hillman, A. M. and Hitt, M. A. (1999). Corporate Political Strategy Formulation: A Model of Approach, Participation, and Strategy Decisions. *Academy of Management Review* 24 (4): 825–842.

- Itami, H. and Roehl, T. W. (1987). *Mobilizing Invisible Assets*. Harvard University Press, Boston, MA.
- Kazanjian, R. K. and Rao, H. (1999). The Creation of Capabilities in New Ventures: A Longitudinal Study. *Organization Studies* 20 (1): 125–142.
- Keim, G. (1981). Foundations of a Political Strategy for Business. *California Management Review* 23 (4): 41–48.
- Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* 3 (3): 383–397.
- Leonard-Barton, D. (1995). *Wellsprings of Knowledge: Building and Sustaining the Source of Innovation*. Harvard Business School Press, Boston, Mass.
- Li, H. and Atuahene-Gima K. (2001). Product Innovation Strategy and Performance of New Technology Ventures in China. *Academy of Management Journal* 44 (6): 1123–1134.
- Lord, D. M. (2003). Constituency Building as the Foundation for Corporate Political Strategy. *Academy of Management Executive* 17 (1): 112–124.
- MacMillan, I. C. and Day, D. L. (1987). Corporate Ventures into Industrial Markets: Dynamics of Aggressive Entry. *Journal of Business Venturing* 2: 29–40.
- McDougall, P. P. (1989). International versus Domestic Entrepreneurship: New Venture Strategic Behavior and Industry Structure. *Journal of Business Venturing* 4: 387–399.
- McDougall, P. P., Robinson, R. B. and Herron, L. (1994). The Effects of Industry Growth and Strategic Breadth on New Venture Performance and Strategy Content. *Strategic Management Journal* 15: 537–554.
- McDougall, P. P., Robinson, R. B. and DeNisi, A. S. (1992). Modeling New Venture Performance: An Analysis of New Venture Strategy, Industry Structure, and Venture Origin. *Journal of Business Venturing* 7: 267–289.
- McGee, J. E., Dowling, M. J. and Megginson, W. L. (1995). Cooperative Strategy and New Venture Performance: The Role of Business Strategy and Management Experience. *Strategic Management Journal* 16: 565–580.
- McGrath, R. G. (1995). Advantage from Adversity: Learning from Disappointment in Internal Corporate Ventures. *Journal of Business Venturing* 10 (2): 121–133.
- Miles, R. E. and Snow, C. (1978). *Organizational Strategy, Structure and Process*. McGraw Hill, New York.
- Miller, A. and Camp, B. (1985). Exploring Determinants of Success in Corporate Ventures. *Journal of Business Venturing* 1: 87–105.
- Miller, D. (2003). An Asymmetry-Based View of Advantage: Towards an Attainable Sustainability. *Strategic Management Journal* 24 (10): 961–976.
- Moorman, C. and Miner, A. S. (1998). Organizational Improvisation and Organizational Memory. *Academy of Management Review* 23 (4): 698–723.
- Moran, P. (1998). Personality Characteristics and Growth-Oriented of the Small Business Owner-Manager. *International Small Business Journal* 16 (3): 17–38.
- Nelson, R. and Winter, S. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, MA.
- Nelson, R. (1991). Why do Firms Differ, and How does it Matter? *Strategic Management Journal* 12 (8): 61–74.
- Park, S. and Bae, Z.-T. (2004). New Venture Strategies in a Developing Country: Identifying a Typology and Examining Growth Patterns through Case Studies. *Journal of Business Venturing* 19: 81–105.
- Penrose, E. (1959). *The Theory of the Growth of the Firm*. Oxford University Press, Oxford.
- Porter, M. E. (1980). *Competitive Strategy*. The Free Press, New York.
- Porter, M. E. (1985). *Competitive Advantage*. The Free Press, New York.



- Reynolds, P. D., Camp, S. M., Bygrave, W. D., Autio, E. and Hay, M. (2001). *Global Entrepreneurship Monitor*. 2001 Executive Report. Kauffman Foundation, Kansas, MO.
- Robinson, K. C. and McDougall, P. (1998). The Impact of Alternative Operationalizations of Industry Structural Elements on Measures of Performance for Entrepreneurial Manufacturing Ventures. *Strategic Management Journal* 19: 1079–1100.
- Robinson, K. C. (1999). An Examination of the Influence of Industry Structure on Eight Alternative Measures of New Venture Performance for High Potential Independent New Ventures. *Journal of Business Venturing* 14: 165–187.
- Romanelli, E. (1989). Environments and Strategies of Organization Start-up: Effects on Early Survival. *Administrative Science Quarterly* 34: 369–387.
- Sanchez, R. (1995). Strategic Flexibility in Product Competition. *Strategic Management Journal* 16: 135–159.
- Sandberg, W. (1986). *New Venture Performance: The Role of Strategy and Industry Structure*. Lexington Books, Lexington, MA.
- Shepherd, D. A. and Shanley, M. (1998). *New Venture Strategy*. Sage Publications Inc., Thousand Oaks, CA.
- Shrader, R. and Simon, M. (1997). Corporate versus Independent New Venture: Resource, Strategy, and Performance Differences. *Journal of Business Venturing* 12: 47–66.
- Simon, M. and Houghton, S. (2002). The Relationship among Biases, Misperceptions and Introducing Pioneering Products: Examining Differences in Venture Decision Contexts. *Entrepreneurship: Theory & Practice* 27 (2): 105–124.
- Stinchcombe, A. L. (1965). Social Structure and Organizations. In *Handbook of Organizations*, J. G. March (Ed.), pp. 142–194. Rand-McNally & Co, Chicago.
- Teece, D. J., Pisano, G. and Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal* 18 (7): 509–533.
- Timmons, J. (2005). *Entrepreneurship*. Price—Babson Symposium, Babson Park, MA.
- Vesper, K. H. (1990). *New Venture Strategies*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ.
- Weidenbaum, M. (1980). Public Policy: No Longer a Spectator Sport for Business. *Journal of Business Strategy* 3 (4): 46–53.
- Wiklund, J. and Shepherd, D. A. (2003). Aspiring for and Achieving Growth: The Moderating Role of Resources and Opportunities. *Journal of Management Studies* 40: 1919–1441.
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. Free Press, New York, NY.
- Winter, S. G. (2003). Understanding Dynamic Capabilities. *Strategic Management Journal* 24 (10): 991–995.
- Zahra, S. A. (1996a). Technology Strategy and New Venture Performance: A Study of Corporate-Sponsored and Independent Biotechnology Ventures. *Journal of Business Venturing* 11 (4): 289–321.
- Zahra, S. A. (1996b). Technology Strategy and Financial Performance: Examining the Moderating Role of the Firm's Competitive Environment. *Journal of Business Venturing* 11: 189–219.
- Zahra, S. A. (2005). A Theory on International New Ventures: A Decade of Research. *Journal of International Business Studies* 36 (1): in press.
- Zahra, S. A. and Bogner, W. (2000). Technology Strategy and Software New Venture Performance: The Moderating Effect of the Competitive Environment. *Journal of Business Venturing* 15 (2): 135–173.
- Zahra, S. A., Ireland, D. R. and Hitt, M. (2000). International Expansion, Technological Learning and New Venture Performance. *Academy of Management Journal* 43: 925–950.
- Zahra, S. A. and George, G. (2002a). Absorptive Capacity: A Review, Reconceptualization, and Extension. *Academy of Management Review* 27: 185–203.
- Zahra, S. A. and George, G. (2002b). International Entrepreneurship: The Current Status of the Field and Future Research Agenda. In *Strategic Entrepreneurship: Creating a New Mindset*, M. A. Hitt, R. D. Ireland and D. L. Sexton (Eds), pp. 255–288. Blackwell, London, UK.

- Zahra, S. A. and Kirchof, B. (2005). Technological Resources and New Venture Growth. *Research in Sociology [Special issue on Entrepreneurship]* 15: 103–124.
- Zahra, S. A., Nash, S. and Bickford, D. (1995). Transforming Technological Pioneering into Competitive Advantage. *Academy of Management Executive* 9 (1): 17–31.
- Zahra, S. A. and Nielsen, A. P. (2002). Sources of Capabilities, Integration and Technology Commercialization. *Strategic Management Journal* 23: 377–398.
- Zahra, S. A. and Pearce, J. (1990). Research Evidence on the Miles-Snow Typology. *Journal of Management* 16: 751–768.

## Stage 2: Aspects of Entry and New Venture Creation

LENA LEE and POH KAM WONG

*National University of Singapore, Entrepreneurship Centre, E3A Level 6,  
10 Kent Ridge Crescent, Singapore 119260*

*E-mail: lenalee@nus.edu.sg*

## 4. Entrepreneurship Education—A Compendium of Related Issues

### 1. INTRODUCTION

A wide range of factors has contributed to the revival of interest in entrepreneurship in the U.S., Europe and many other countries where we define entrepreneurship as the emergence and growth of new businesses (Rosa et al., 1996). In recent years, many industrialized countries have suffered from economic recessions and high unemployment rates. Given the prevailing economic conditions, policy makers worldwide have now begun to recognize the instrumental role of entrepreneurship for economic growth. New and growing businesses are seen as a solution to rising unemployment rates and as a major catalyst to national economic prosperity (Acs et al., 1999; Bruyat and Julien, 2000).

As a result of the proliferating worldwide emphasis on entrepreneurship as the catalyst for economic development and job creation, policy makers have developed a wide array of measures to support entrepreneurship (Gnyawali and Fogel, 1994; Maillat, 1998). Key among these is the call for academic institutions, such as universities, to contribute through appropriate educational programs, that is, entrepreneurship education (Laukkanen, 2000). There appears to be a consensus that entrepreneurship education and training has a major role to play in the economic development of a country (Gibb, 1996). The nature, relevance and appropriateness of entrepreneurship education have been subject to increasing scrutiny since the late 1960s (Vesper, 1985) and it has been touted as an effective means of entrepreneurial learning (Johannisson, 1991). Indeed, entrepreneurship education has become an obvious complement to venture

capital and incubators as tools in propelling economic advancement (McMullan and Long, 1987).

While there is clearly a rise in entrepreneurial spirit and, hence, an increase in entrepreneurship education worldwide,<sup>1</sup> information about the roles of entrepreneurship education remains largely scattered and sporadic. This chapter aims to consolidate and synthesize the issues surrounding university entrepreneurship education, such as whether or not entrepreneurship can be taught, the structure, effectiveness and potential growth of entrepreneurship courses and ethical issues, as well as new perspectives on entrepreneurship education.

## 2. ENTREPRENEURSHIP EDUCATION—AN OVERVIEW

Interest in entrepreneurship education grew rapidly in the late 1950s and early 1960s when some of the most well-known studies in entrepreneurship like McClelland's *The Achieving Society* (1961) and Collins, Moore, and Umwalla's *The Enterprising Man* (1964) were published. Over the years, entrepreneurship education has climbed the ranks in the business domain and was positioned sixth in importance out of the 60 recommendations on the solutions to the major problems facing small businesses presented at the White House Conference on Small Business (Solomon and Fernald, 1991).

There are various definitive labels used to explain entrepreneurship education. The term "entrepreneurship education" is commonly used in the U.S. and Canada but was less commonly applied in Europe in the early 1980s (Gibb, 1993). The preferred term within the U.K. and Irish contexts was "enterprise education," but by the early 1990s, the concept of "enterprise" gradually converged to "entrepreneurship" (Gibb, 1993). Essentially, Gibb made a clear distinction between "enterprise" and "entrepreneurship" education with the former focusing on the advancement of personal enterprising attributes and attitudes that prepare the individual for self-employment, while the latter relates to the development of functional management skills and abilities that train the individual to start, manage and develop a business (Gibb and Nelson, 1996, p. 98). Despite this distinction, the ultimate aim of both enterprise and entrepreneurship education is to encourage independent business creation.

### 2.1. *The Dichotomy of Entrepreneurship: Nature vs. Nurture*

In the midst of continuous growth in the number of universities offering entrepreneurship courses, opinions abound on the issue of whether entrepreneurship can be taught and anecdotes about whether entrepreneurs are born or bred fill discussions in international journals and conferences. Commentators

such as Rae and Carswell (2001) and Shepherd and Douglas (1997) argued that there is a distinction between the teachable and the nonteachable elements of entrepreneurship. The key to a successful entrepreneurship education is to find the most effective way to manage the teachable skills and identify the best match between student needs and teaching techniques (Katz, 1991). This concurs with the findings of an earlier study conducted by Vesper (1971) who highlighted that the debate over not how entrepreneurship can be taught but how it can best be taught.

As discussed by Jack and Anderson (1998), the teaching of entrepreneurship is both a science and an art where the former relates to the functional skills required for business start-up (an area which appears to be teachable) while the latter refers to the creative aspects of entrepreneurship, which are not explicitly teachable. There is a unanimous agreement among entrepreneurship educators that there needs to be a shift of emphasis from the scientific to the artistic and creative teaching of entrepreneurship (Shepherd and Douglas, 1997). Although the focus of most entrepreneurship courses and training lies in the scientific dimension of entrepreneurship, it has been acknowledged that entrepreneurship education helps ignite the artistic, creative and perceptual aspects of entrepreneurship (Shepherd and Douglas, 1997).

Indeed, recent evidence in the literature indicates that entrepreneurship education has a positive impact on perceptual factors such as self-efficacy (Cox et al., 2002/2003). The authors compared students who had not yet begun an entrepreneurship course (pre-course group) with those who had completed the course (post-course group) and found that the post-course group had significantly higher self-efficacy than the pre-course group. Similarly, in a pre-test/post-test study, Peterman and Kennedy (2003) observed that participants reported significantly higher perceptions of both desirability and feasibility after completing the Young Achievement Australia (YAA) entrepreneurship program.

From a different perspective, Dana (2001) posited that the question of whether entrepreneurship can be taught depends on the fundamental definition of entrepreneurship. He discussed both the Schumpeterian and Austrian definitions of entrepreneurship and argued that it is possible to train potential entrepreneurs to identify opportunities but difficult to teach them the art of creating opportunities. Fundamentally, Dana argued that Kirznerian entrepreneurship (opportunity identification) is teachable but not Schumpeterian entrepreneurship (opportunity creation). Further, Saks and Gaglio (2002) added that while it is possible to teach participants of entrepreneurship programs to evaluate opportunities, the innate ability to recognize opportunities remains virtually nonteachable. As Saeed (1996) asserted, entrepreneurship education can only demonstrate the process involved in being successful, but cannot create an entrepreneur as the individual is ultimately responsible for his/her own success.

Collectively, evidence in the existing literature delineates that only specific aspects of entrepreneurship are explicitly teachable, and it may be necessary to teach people how to be entrepreneurs (Garavan and O’Cinneide, 1994). As Professor Howard Stevenson of Harvard University eloquently put it: “You cannot teach someone to become a Bill Gates, (neither can you) teach someone to compose like Beethoven. But you can teach someone the notes and scales, give them the tools they need to become a composer. And you can teach the tools people need to be entrepreneurs.” Therefore, the issues surrounding the provision of entrepreneurship education merits further attention and are explored further in the following sections.

### 3. STRUCTURE OF ENTREPRENEURSHIP COURSES

#### *3.1. Objectives of Entrepreneurship Education*

Generally, entrepreneurship education aims to increase the awareness of entrepreneurship as a career option, and enhance the understanding of the process involved in initiating and managing a new business enterprise (Hills, 1988; Donckels, 1991). Following Interman’s (1992) detailed typology of entrepreneurship, there are four objectives of entrepreneurship programs: entrepreneurship awareness, business creation, small business development, and training of trainers.<sup>2</sup> Similarly, Jamieson (1984) suggested that entrepreneurship education provides three different classes of training: education “about” enterprise (i.e., entrepreneurship awareness), education “for” enterprise (i.e., preparation of aspiring entrepreneurs for business creation) and education “in” enterprise (i.e., training for the growth and development of established entrepreneurs). A recent study by Parker (2005) stressed the importance of education and training for existing entrepreneurs as it was found that entrepreneurs learned rather slowly and they tend to rely disproportionately on prior beliefs and past experiences.

In contrast, commentators such as Johannisson (1991) posited that entrepreneurship education has five learning objectives in that participants of entrepreneurship programs will develop the know why (developing the right attitudes and motivation for start-up); know how (acquiring the technical abilities and skills needed to develop a business); know who (fostering networks and contacts for entrepreneurial ventures); know when (achieving the sharp intuition to act at the correct moment); and know what (attaining the knowledge base and information for new venture development) aspects of entrepreneurial learning.

In a broader context, the four main objectives of entrepreneurship education appear to be i) prepare participants for career success; ii) increase their capacity for future learning; iii) realize participants’ personal fulfill-

ment; and iv) contribute to society (Sexton and Kasarda, 1992). Garavan and O’Cinneide (1994) extended this characterization by advancing that the goal of entrepreneurship is to effectively foster all these objectives for the creation of new businesses. Given the high investment of resources in entrepreneurship education,<sup>3</sup> a general consensus emerges in the literature that the primary aim of these programs is the promotion of the successful formation of new ventures (Curran and Stanworth, 1989). Supporting this assertion, McMullan and Long (1987) argued that unlike other university degrees, the success of entrepreneurship programs cannot be evaluated by the number of students graduated but more appropriately measured by the socioeconomic impact they produce in the businesses they create. Issues such as the number of companies created, the number of jobs created, the types of companies formed and the growth potential of the companies are essential for economic growth (Sternberg and Wennekers, 2005; Wong et al., 2005).

The view that entrepreneurship education and, hence, entrepreneurship courses need to be differentiated by stages of venture development rather than by functional expertise has been acknowledged for some time (McMullan and Long, 1987).<sup>4</sup> Specifically, the implications that the learning needs of individuals vary according to the stages of development such as awareness, pre-start-up, start-up, growth and maturity. According to McMullan and Long, the typical skill required at a particular stage is as follows: Opportunity identification at the awareness stage, market feasibility at pre-start-up and new venture planning during start-up.

Although there are no definitive objectives of entrepreneurship education, the aims of such programs and trainings are probably best summarized as i) identifying and preparing potential entrepreneurs for start-ups; ii) enabling participants to prepare business plans for new venture; iii) focusing on issues that are critical to the implementation of entrepreneurial projects such as market research, business financing and legal issues; and iv) enabling the development of autonomous and risk-taking behavior (Garavan and O’Cinneide, 1994).

### *3.2. Content of Entrepreneurship Education Courses*

The philosophical underpinning of entrepreneurship courses is that the participants of these courses can influence the external environment and that they are not bounded by the intricacies in the environment. This philosophical perspective is closely related to the locus of control theory that elucidates the positive relationship between internal locus of control and entrepreneurial start-ups (Hansemark, 2003). Studies by Hansemark (1998) and Jennings and Zeithamil (1983) reported that participation in an entrepreneurship program increases internal locus of control. It is expected that individuals are able to apply the skills and knowledge that they have acquired through entrepreneurship



education and training to venture-related decision making. Some of the typical skills required for start-ups are knowledge on how to raise finance, the legal and tax framework, marketing and recruitment (Garavan and O'Conneide, 1994) resulting in the development of more practical-based entrepreneurship programs at the expense of conceptual development (Sexton and Bowman, 1984).

In response to entrepreneurs' demands for entrepreneurship education, a new age of learning that consists of outreach programs is offered to existing entrepreneurs instead of the usual pool of students (McMullan et al., 1986). Basically, the common elements in an entrepreneurship course include lectures, venture plan writing, entrepreneurial speakers, business cases and more recently, the use of live video of entrepreneurs featured in cases (Gartner and Vesper, 1994). In an attempt to advocate a framework and methodology for entrepreneurship education, Knight (1987) suggested that the following elements be included in entrepreneurship programs: opportunity identification, strategy development, resource acquisition and implementation.

Supporting McMullan and Long's (1987) proposition that entrepreneurship education should be structured based on the different skills needed at various stages of the firm's development, Gartner and Vesper (1994) observed that the skills and knowledge required to understand business entry (entrepreneurship) differ from the skills and knowledge required to comprehend the operations of an ongoing business (business management). Gibb (1993) distinguished the learning focus of business school from entrepreneurship education. He argued that some entrepreneurship programs employ the curriculum of business schools that is not compatible in an entrepreneurial situation.<sup>5</sup>

The values and abilities emphasized by business schools may actually inhibit entrepreneurial spirit. As noted by Kao (1994), the management model of teaching does not apply to entrepreneurship; hence, distinctive curricula and training programs are needed for entrepreneurship education. However, Zeithaml and Rice (1987) cautioned that, although education for entrepreneurship and education for small business management are not the same thing, the two terms are so closely associated that it is almost impossible to study one without considering the other.<sup>6</sup> Entrepreneurship courses stress the equivocal elements of start-ups such as the development of new organizations, new products and new markets while business management courses emphasize the knowledge and skills required for business practices.

In terms of the division between undergraduate and graduate courses in entrepreneurship, more courses are offered at the undergraduate level in universities (Gartner and Vesper, 1994). Brush et al. (2003), in their research on entrepreneurship education at different levels, found that deans of business schools were increasingly placing more importance on entrepreneurship at the undergraduate than at the graduate levels. From the contents aspect, Zeithaml and Rice's (1987) survey of schools of varying sizes and their offerings of

entrepreneurship courses at the graduate and undergraduate levels suggested that the structure of entrepreneurship education programs are quite similar throughout the United States at both levels (i.e., each school offers one general course usually aimed at the development of a new business, followed by consulting experience with a small business that is facing difficulties). In terms of duration, most of these entrepreneurship programs last as short as a few days (Curran and Stanworth, 1989), while others range from 25 days to 12 months (Garavan and O’Cinneide, 1994).

Robinson and Haynes (1991) introduced the terms “depth” and “breadth” into the context of entrepreneurship education programs. Depth relates to the quality of program, while breadth refers to the number of entrepreneurship programs available. The authors proposed that the higher the quality of the program, the greater the commitment to, and formalization of academic programs, the more will be the institutional resources committed, the higher will be the financial aid and the greater will be the number of extracurricular organizations (clubs, societies) available. In their survey of 215 colleges and universities, Robinson and Haynes reported that 81.5% of the schools they studied had at least one undergraduate course and 54.2% offered one graduate course. Additionally, 61.1% of the universities had a formal organization, for example, centers/departments established for entrepreneurship education.

The challenge for entrepreneurship educators is to enhance the quality of existing programs in existing institutions, that is, improving the depth of the field and not merely extending existing entrepreneurship courses and programs to other institutions, that is, developing the breadth of the field (Robinson and Haynes, 1991). However, the major stumbling block for such development has been the lack of solid theoretical bases upon which pedagogical models and methods are built. Thus, the authors affirmed the need for institutions of higher learning to develop graduate doctoral programs that prepare future faculty in entrepreneurship education.

Plaschka and Welsch (1990) found some commentators in the field who believe that it is possible to map out a framework of entrepreneurship programs that aims at developing a competent curriculum for entrepreneurship education. With this objective in mind, Plaschka and Welsch proposed two frameworks. The first framework consists of two dimensions: the number of entrepreneurship courses that are offered and the degree of integration. The number of courses range from a single course to multiple courses, while the degree of integration represents the level of acceptance and support from a variety of different groups. Support can be sought from intra-university groups such as other faculty members, inter-university groups such as alumni and entrepreneurs and complementary entrepreneurship activities such as entrepreneurship clubs or organizations. This framework is a matrix that extends from “unsupported

isolated course” to “integrated program.”<sup>7</sup> When more and more courses are offered to participants, the ideal path will be for education providers to adopt the “integrated program” strategy where multiple courses are integrated both with one another and with the curriculum.

The second framework is based on two paths: transition stages and functional fields. The transition stages are inception, survival, growth, expansion and maturity, while functional fields refer to the different disciplines of the entrepreneurship curriculum such as marketing, management and finance. Entrepreneurship educators could conduct courses within a single discipline like marketing, focusing on start-ups (“Unidisciplinary Approach”) or introduce multiple disciplines in the entrepreneurship curriculum, focusing on mature firms (“Interdisciplinary Program”).<sup>8</sup> This framework suggests that there is no best strategy for curriculum development as the choice of design depends on the stage of the firm’s development and the focus of the curriculum (uni-discipline vs. multi-discipline).

In a recent study, B  chard and Toulouse (1998) provided a comprehensive review of how entrepreneurship programs are developed. Based on their report, they proposed that the contents of entrepreneurship programs can be planned from four perspectives. First, program contents can be developed from the perspective of the educators, where the curriculum is defined based on the expertise of the educators. Second, entrepreneurship program can be established based on the students’ needs and requirements. This approach takes into account the learning requirements of each individual. Third, the entrepreneurship syllabus can be analyzed from the viewpoint of those who design them. This view considers the key learning/teaching objectives as the anchor of entrepreneurship curriculum. Fourth, the evaluators of the programs themselves can influence the curriculum. This approach allows evaluators to make adjustments to the program contents according to the pre-set criteria for program quality and effectiveness. Depending on the host institution and level of priority, there is no one best perspective for program development.<sup>9</sup>

From the review presented, it is evident that despite the increase in entrepreneurship programs, there is still no generally accepted curriculum for aspiring entrepreneurs to follow (Plaschka and Welsch, 1990; Koch, 2002/2003). Many programs evolve on a trial-and-error basis, depending on the types of entrepreneurial projects undertaken and on the feedback of students who have enrolled on the courses. However, the authors cautioned that these are not necessarily poor approaches to program development as long as there is a mechanism that systematically documents the feedback. Moreover, given the changing nature and relatively young discipline of entrepreneurship, Saks and Gaglio (2002) noted that it will be impossible to find unanimous agreement on the teaching content of entrepreneurship education.

### *3.3. Approaches to Entrepreneurship Education*

The considerable variety of entrepreneurial programs offered in the market is also found in the variety of learning methods employed in entrepreneurship education and training. From a broad perspective, the approaches to entrepreneurship education can be classified into four categories: the “old war stories” approach, the “case study” approach, the “planning” approach and the “generic action” approach (Shepherd and Douglas, 1997). Each of these approaches differs from each other in terms of its focus and purpose. The “old war stories” approach attempts to motivate aspiring entrepreneurs by relaying a series of successful entrepreneurship stories and revealing how those individuals became successful entrepreneurs. The “case study” approach uses cases of existing companies to analyze the mechanics of the entrepreneurial process and to elicit students’ proposed solutions to the companies’ problems. The “planning” approach usually takes the form of a business plan that consists of detailed objectives, budgets and programs, while the “generic action” approach emphasizes the formulation of optimal entrepreneurial actions based on existing market forces.

On the other hand, from a micro perspective, the study by Solomon et al. (2002) highlighted that the most popular teaching methods in entrepreneurship education are creation of business plans, case studies and lectures. Some commentators, like McMullan and Long (1987), are of the opinion that entrepreneurship education should be creatively grounded and that students should be exposed to problem solving and taught strategies to deal with ambiguous and complex situations. Apart from this, students should also be exposed to substantial hands-on working experience with community ventures. Stumpf et al. (1991) supported this view when they highlighted the applicability of behavioral simulations in entrepreneurship programs. Hills (1988), in his survey of 15 leading educators in the field, found that courses had been created around the production of business plans with an emphasis on market feasibility analysis. Additionally, he reported that educators preferred to develop courses around business life cycles, that is, start-up, growing firms and established firms. In other studies, McMullan and Boberg (1991) and Preshing (1991) observed that students generally favored the project method as compared to the case method of teaching entrepreneurship.

Further to the recommendation of approaches in entrepreneurship education, Carsrud (1991) postulated that apart from full-time academia, other components of the entrepreneurship infrastructure such as endowed faculty, research centers, professional organizations, journals and the mass media provided significant support for entrepreneurship education. Sandberg and Gatewood (1991) highlighted the pivotal role of entrepreneurship centers as the intermediary between faculty and business community. In essence, the agendas and constituency orientations of entrepreneurship centers facilitated the

teaching of entrepreneurship by bridging the gap between business schools and the community.

Thus, while there is a substantial array of teaching methods employed in entrepreneurship education, there is little intellectual cohesion among these efforts and the literature seems to indicate that there is no one best pedagogical approach for teaching entrepreneurship (Garavan and O’Cinneide, 1994; Kolb et al., 1974). Nevertheless, the “learning” approach proposed by Shepherd and Douglas (1997) has been touted as an effective path toward developing the entrepreneurial spirit. This approach requires the shift in emphasis from teaching to learning and it recognizes the importance of learning through hands-on and active participation in a real life entrepreneurial environment where constructive feedback from an expert is provided.

### *3.4. Profile of Participants of Entrepreneurship Education Programs*

There is great diversity among the participants of entrepreneurship programs, and various categories have been proposed in the literature. Generally, a high percentage of students are practitioners who choose to embark on an entrepreneurial program to acquire new knowledge and skills which they can assimilate into their businesses (McMullan and Long, 1987). Garavan and O’Cinneide (1994) identified three categories of participants: first, participants who had no project idea for starting a business but who within a specified period of starting the program would find one; second, participants who already had a concrete idea of starting a business; and third, participants who had only a basic and tentative idea of starting a business. In stark contrast, Birley (1984) classified entrepreneurship into three sets of potential customers: conventional full-time students; entrepreneurs (those who are just starting businesses and those who are managing ongoing businesses); and third, business advisors such as bankers, accountants and government policy makers.

From a planning viewpoint, Hynes (1996) suggested that entrepreneurship education should be incorporated into the nonbusiness disciplines of engineering and science where product ideas emerge but are often ignored because students are not sufficiently educated in the knowledge and skills required for start-ups. In light of this, there is a growing trend toward university-wide programs in entrepreneurship education (Streeter et al., 2002). According to Streeter and his associates, the core objective of university-wide programs is to extend the opportunity of entrepreneurship education to all students in the university regardless of their faculty or subject major. As entrepreneurship education transcends all fields of studies, it is no longer associated exclusively with business graduates. In response to the promising appeal of entrepreneurship education across faculties, business schools have joined forces with other departments within the university to offer students customized entrepreneur-

ship courses where the curriculum addresses the integration of the complex processes of entrepreneurship with respect to the specialized technical areas of their respective field of studies (Streeter et al., 2002).

Further evidence of the value of entrepreneurship education for non-business disciplines of engineering and science is highlighted in a recent study of high-tech entrepreneurs in the U.S. where the lack of knowledge and understanding about starting a business was perceived as a major obstacle to high-tech entrepreneurship (Kourilsky and Walstad, 2002).<sup>10</sup> In a broader context, Brush (1992) and Krueger (1993) proposed that entrepreneurship education should also be targeted at students from families with entrepreneurs as these individuals generally tend to have a more positive attitude toward this type of education.

In recent years, both researchers and policy makers, particularly in the U.K., have hailed the need for entrepreneurship educators to focus on students at the pre-university level, that is, secondary school. Commentators like Chamard (1989) and Singh (1990) believe that existing secondary school systems inhibit entrepreneurship and do not foster the development of an enterprise culture among the students. Similarly, Filion (1994) and Gasse (1985) emphasized the importance of identifying and cultivating entrepreneurial potential at the secondary school level when the individuals' career options are still open.

The key for entrepreneurship education providers is to understand that, given the wide variety of entrepreneurship programs being offered, it is not surprising that different types of programs will attract different groups of participants (Jack and Anderson, 1998). Not all participants of entrepreneurship education intend to start a new business; some simply embark on a program to enhance their knowledge of a field that is presently very popular and rapidly growing (Block and Stumpf, 1992; Jack and Anderson, 1998). The challenge for entrepreneurship educators is to target their courses and training programs to all interested parties, including those who have no direct intention to become entrepreneurs because these noninterested groups may very well provide the supporting base to existing and potential entrepreneurs.

However, it is important that the entrepreneurship curricula match the needs of different participants. Past studies have shown that entrepreneurs are often reluctant to participate in the courses offered by local colleges and university because they perceived these courses to be too theoretical and "academically oriented" (McCarthy et al., 1997). For example, in response to the possible mismatch between the expectations of undergraduate students and the business community, McCarthy et al. proposed an entrepreneurship program that enhances the real-life practical experiences of these students while providing assistance to businesses that are struggling with post start-up problems. Their proposed model aimed at benefiting both the undergraduates and the business community.

### 3.5. *Deliverers of Entrepreneurship Courses*

The literature calls for a balance between academics and practitioners, known as teamwork teaching in the delivery of entrepreneurship education (McMullan and Long, 1987).<sup>11</sup> Academics usually contribute and provide evidence based on theoretically grounded studies while practitioners teach by providing practical examples of how to make things happen. Practitioners who teach entrepreneurship are given a job title that reflects this nonfaculty status, for example, Adjunct Faculty. There is a general consensus in the field that adjunct faculties are “useful additions” to the teaching team, however, they are not builders of intellectual capital, which is critical for long-term legitimacy (Hills, 1988). Entrepreneurship education may also entail cooperative teaching from faculties across a number of different schools such as management, engineering, law and computer science. Endowed positions or chairs in entrepreneurship are common among members of the entrepreneurship teaching team (Katz, 1991). Based on Katz’s study, the first endowed position in entrepreneurship was in 1963 at Georgia State University, and the second one was in 1975. By 1985, there were a total of 25 endowed positions in the U.S., and the number has continued rising over the years, reaching a total of 102, 123, 271 and 406 positions in 1990, 1994, 1999 and 2003 respectively.<sup>12</sup> Endowed positions outside of the U.S. grew from four to 34 in 1991, and 158 in 1999 and 2003, respectively.

A major weakness in the supply of entrepreneurship educators is the lack of doctoral programs in entrepreneurship to send trained academics into the career pipeline (Saks and Gaglio, 2002). Even with existing entrepreneurship doctorates, these individuals usually come from other disciplines like organizational behavior, marketing and finance, producing a wide range of backgrounds among entrepreneurship researchers. The cross-disciplinary background of these researchers have contributed to the plethora of entrepreneurship studies that draw on various theories from the fields of psychology (e.g., Krueger and Dickson, 1994), economics (e.g., Evans and Leighton, 1990), strategy (e.g., Hitt et al., 2001) and organizational behavior (e.g., Hult et al., 2003). However, the diverse backgrounds of these scholars have contributed to entrepreneurship education’s struggle in establishing itself as a legitimate field distinct from other disciplines (Brush et al., 2003). What accounts for the lack of specific entrepreneurial doctoral programs? The answer probably lies in the inherent difficulty of developing an entrepreneurship program that balances both research and practical business. The rapid growth of entrepreneurship education at the undergraduate level has far outstripped the development of postgraduate courses, particularly doctorate offerings in entrepreneurship, resulting in an under-supply of doctoral-trained entrepreneurship faculty to develop the field’s scholarly research and knowledge (Brush et al., 2003).

#### 4. EFFECTIVENESS OF ENTREPRENEURSHIP EDUCATION

Many researchers including Block and Stumpf (1992) and Curran and Stanworth (1989) have identified the need for evaluating entrepreneurship education and training programs. In the extant literature, there are numerous studies that attempt to measure the effectiveness of entrepreneurship education and training. Yet, implementing an effective research design to isolate the effects of different programs across universities is a monumental task. The vast majority of studies that attempt to examine the link between entrepreneurship education and new venture creation suffer from intrinsic procedural and methodological limitations (Curran and Storey, 2002; Gorman and Hanlon, 1997).<sup>13</sup> Longitudinal research designs, using control groups to compare participants with individuals who did not have entrepreneurial educational experience, are needed to examine the lasting effects of entrepreneurship education and training interventions. As Gorman and Hanlon (1997: 71) asserted, “since the cumulative impact of repeated exposure to education for entrepreneurship should be expected to have a much greater impact on attitudes and propensity, a difficult but important challenge for researchers will be to measure the overall effectiveness of these programs.”

There is unequivocal consensus among researchers that one of the primary economic measures of entrepreneurship program effectiveness is the number of new businesses started (McMullan et al., 2001). The literature provides evidence of the positive relationship between entrepreneurship education and the number of venture start-ups. Individuals who have attended entrepreneurship courses have a higher tendency to start their own businesses at some point in their career than those who attended other courses (Charney and Libecap, 2000<sup>14</sup>; McMullan and Gillin, 1998).<sup>15</sup> Clouse (1990),<sup>16</sup> Garnier and Gasse (1990)<sup>17</sup> and Garnier et al. (1991)<sup>18</sup> provided additional evidence that participation in an entrepreneurship program has a positive impact on one’s decision to start a new venture. In a similar vein, Price and Monroe (1992)<sup>19</sup> found that entrepreneurship training has a positive relationship with venture growth and development.

McMullan et al. (1985) measured venture creation activities of students taking three or more new venture development courses at the MBA level of University of Calgary and found a relatively high start-up rate among the graduates, that is, 14% of the graduates started businesses. Similarly, Brown et al. (1987) surveyed participants of an entrepreneurship program called “Your Future in Business” that aims to address the education needs of potential and existing entrepreneurs, and they concluded that a significant number of new ventures were created by the graduates of the program. Furthermore, education and training of entrepreneurs have been repeatedly cited as an effective way to reduce small business failure (Carrier, 1999).<sup>20</sup>



The contribution of entrepreneurship education in society is well documented in the literature. As noted by Galloway and Brown (2002), in addition to developing skills for business start-up and ownership, entrepreneurship education makes a significant contribution in terms of the quality of graduate start-ups and it influences general attitudes to entrepreneurship in the long term. In Galloway and Brown's view, entrepreneurship education represents a positive motivation in terms of promoting entrepreneurship as a respectable and valuable career option.

Entrepreneurship education inevitably influences the population's attitudes toward entrepreneurship and assists in the creation and maintenance of an enterprising culture. In the long run, it helps build a risk tolerant and entrepreneurial society. Entrepreneurship courses also prepare participants for intrapreneurial challenges in large corporations where skills such as creativity, innovation and proactiveness are essential. As an integral component in the venture support system, entrepreneurship education complements incubators, science parks and venture capital operations in backing various actors of the entrepreneurial economy. These actors may use the knowledge and know how acquired from the entrepreneurship courses to grow existing businesses. In addition, these courses may also serve as entrepreneurial networking platforms for the participants. Hence, the evaluation of the effectiveness of entrepreneurship education goes beyond traditional business start-up measures.

## 5. POTENTIAL GROWTH OF ENTREPRENEURSHIP EDUCATION

The growth in entrepreneurship education is evident by the avalanche of entrepreneurship centers, chairs, conferences, journals and programs (Plaschka and Welsch, 1990). In recent years, entrepreneurship education has been associated with rising professional associations that operate through a network of formal and informal groups. At the conferences of these associations, theoretician-researchers socialize with practitioners, junior members meet with senior authors and professors, and colleagues join together to share recent information, developments and innovations in the field.<sup>21</sup> The growth of entrepreneurship education goes well beyond the U.S. As reported in Brockhaus's (1991) study, entrepreneurship is recognized as a major tool for economic development worldwide, beyond both the U.S. and Europe. Universities worldwide (including former communist countries) not only offer entrepreneurship courses but also conduct a consistent stream of practical and theoretical research on entrepreneurship issues.

The prevailing questions in the field are what the future of entrepreneurship education is and whether it can achieve legitimacy like other business disciplines. The answers are presently unknown but there are many promising signs

indeed; key among these are the development of entrepreneurship programs and centers, the increase in publication outlets and tightening of quality controls on published research, the growth of professional organizations, the proliferation of sources of funding and the rise of endowed positions.<sup>22</sup> Evidence indicates that entrepreneurship education in colleges and universities is spreading rapidly and steadily over the years, and given its positive contribution to the economy, it is not likely that demand for entrepreneurship education will dissipate in the coming years.

The field of entrepreneurship has gained considerable momentum over the years and has achieved the fastest growth in the United States (Dana, 1992). Dana also found that in 1970, 25 institutions of higher learning in the U.S. offered entrepreneurship courses; in 1985, more than 200 institutions joined in the entrepreneurship marathon; and by 1992, entrepreneurship was taught in more than 500 learning establishments. Unfortunately, no studies in the twenty first century have documented the progress of entrepreneurship courses at institutions of higher learning in the U.S. and worldwide, but the founding of the Roundtable on Entrepreneurship Education (REE) USA in 1998 by Stanford University provides a positive indication that entrepreneurship education has permeated many business schools in the U.S.<sup>23</sup> Following the success of REE USA in 1998, the international conference for entrepreneurship educators was extended to other parts of the world, namely REE Europe in 2001, REE Latin America in 2003 and REE Asia in 2004. At all of these REE conferences, a strong participation from the faculties of business and engineering belonging to leading universities was evident. These experts converge to share their entrepreneurial programs' success and failure stories, and to learn from their contemporaries ways to strengthen their entrepreneurship course offerings and training.<sup>24</sup>

As discussed in Section 3.2, the quality of entrepreneurship education is reflected in the depth of its programs, and entrepreneurial programs can be assessed by a number of criteria such as the number of graduate or undergraduate courses offered; the level of commitment to and formalization of the program; the amount of institutional resources available in the form of faculty and staff dedicated to the program, the availability of financial aid for students; and the presence of extracurricular organizations in the form of clubs, societies; and special interest groups supported by the program (Robinson and Haynes, 1991). Overall, evidence in the extant literature indicates that entrepreneurship education has achieved a considerable level of depth in its programs and has indeed met its goal of quality education and training for aspiring and present entrepreneurs. Nevertheless, there are further opportunities for future entrepreneurship education to expand its breadth and depth beyond existing levels.

## 6. ETHICAL CONSIDERATIONS IN ENTREPRENEURSHIP EDUCATION

Despite the intense publicity of entrepreneurship education as an agent for new venture creation, Laukannen (2000) questioned the ethical and moral aspects associated with this publicity. With the tremendous growth of entrepreneurship programs worldwide, individuals both old and young have been inspired and, to some extent, exhorted to start their own businesses. Some people are exposed to entrepreneurship sentiments at a very young age through both the primary and secondary school systems; and some children are conditioned to think that self-employment is the ultimate path for a successful and respectable career.

However, Laukannen provided an intriguing perspective on the ethical issues involved in the provision of entrepreneurship education. He offered an alternative view that society in general is implicitly pressured to take the entrepreneurial plunge notwithstanding the inherent risk and uncertainty involved. Given the “pressure” these people are exposed to, in lieu of the intense interest in entrepreneurial activity, there is a possibility that they may be strongly encouraged under the pretext of economic development. Furthermore, the teaching style of entrepreneurship educators, which appears to be more aggressive than the usual impartial manner of other academicians, seems unorthodox and unneeded in academia. Generally, entrepreneurship educators have a greater expectation of their students to embark on an entrepreneurial career upon completion of the course. Unlike graduates in science or engineering, who may not assume roles associated with their training and education, participants of entrepreneurship courses are sometimes expected to advance their career in self-employment activities.

Entrepreneurship education also tends to overemphasize the contribution of the individual in the creation of new ventures while underplaying the role of teams and existing businesses in the spawning of new businesses. According to Laukannen (2000), most entrepreneurship courses are designed to cover a wide range of business contexts and industries which might be too generalized for the majority of students, particularly those with limited working experience to leverage on for starting new businesses.

However, from a positive perspective, Dyer (1994) proposed that specialized courses and training in entrepreneurship may enhance the individual’s confidence in starting new businesses. Indeed, in a seminal paper, Krueger and Brazeal (1994) highlighted that entrepreneurship education improves perceived feasibility for entrepreneurship by increasing the knowledge of students and promoting self-efficacy among them. Walstad and Kourilsky (1998) provided additional support that entrepreneurship education and training improved desir-

ability for entrepreneurship by convincing students that it is a highly regarded and socially acceptable career in society.

## 7. NEW PERSPECTIVES ON ENTREPRENEURSHIP EDUCATION

In line with the critical role entrepreneurship education serves in the venture-creation system, numerous studies have been carried out to investigate the effects of entrepreneurial education on entrepreneurship. These investigations encompass research areas such as the determination of the positive impact of entrepreneurship education on the decision to start a new venture (Charney and Libecap, 2000;<sup>25</sup> Garnier et al., 1991<sup>26</sup>), the identification of a positive relationship between entrepreneurship education and entrepreneurial success (Ronstadt, 1985; Sexton and Upton, 1987); the investigation of a positive relationship between entrepreneurship education and economic development (McMullan and Long, 1987); and the examination of people's beliefs that entrepreneurship education promotes positive entrepreneurial attitudes (Donckels, 1991; Kantor, 1988).

However, with the exception of a research by Lee and Wong (2003), there has been a lack of research, if any, on the impact of attitude toward entrepreneurship education (hereinafter referred to as AEE) on entrepreneurship. Lee and Wong (2003) argued that it is important to examine AEE because the application of entrepreneurship education alone to explain the entrepreneurial phenomenon may not be sufficient. The authors explained that AEE is critical as there can be a wide array of entrepreneurship courses and training available as part of the venture support system, but if the target group does not perceive that entrepreneurship education will assist them in new venture creation and/or the process of managing a venture, the existence of entrepreneurship education might be redundant as a whole.

Despite the existence of numerous studies that examined the relationship between entrepreneurship education and new business founding, very little is known about the effects of attitudes toward entrepreneurship education on new venture creation. Limited studies in the past have indicated that small business owners possess negative attitudes toward formal education and training (Kailer, 1990) and that most small businesses are prejudiced against participating in formal training (Stanworth and Gray, 1992). However, Marlow (1992) found that minority group entrepreneurs and owners of small firms in the West Midlands, U.K., had considerable interest in formal training.

Although Lee and Wong (2003) found a positive relationship between attitudes toward entrepreneurship education and new venture creation, they did not report the direction of causal link between the two variables. The authors called for future research to establish the causal relationship between AEE and

new venture creation because, according to them, different causal links will yield different implications for academia and policy makers in their pursuit of promoting entrepreneurship.<sup>27</sup>

Entrepreneurship education is also conceived as an important motivating tool for special groups such as women (Price and Monroe, 1992) Blacks (Walstad and Kourilsky, 1998) and ecopreneurs (Anderson and Leal, 1997; Schuyler, 1998). There is a focus on entrepreneurship education for women because they have consistently proven themselves to be a significant driving force in the economy by establishing and expanding businesses at a tremendous pace. The Centre for Women's Business Research (2002) reported that one in 18 women in the U.S. is a business owner, and thus, emphasized the need to provide education and training by extending practical and professional assistance to women in the entrepreneurial community. With five inspiring education and support programs, that is, the Women's Business Enterprise (WBE) certification from the Women's Business Development Center (WBDC), Women's President Organization's (WPO) Minnesota Chapter, Smart Growth Program, Mentoring for Women Business Owners, and Entrepreneurship Minor courses, the Center for Women Entrepreneurs and Entrepreneurship Education at the Metropolitan State University in the U.S. is an active provider of entrepreneurship education and training for women.<sup>28</sup>

In the case of Blacks, evidence in the literature show that the number of black-owned firms is relatively small, accounting for 3.6% of all firms, 1% of total sales and receipts and generating \$52,000 in sales and receipts, compared with an average of \$196,000 overall for businesses. Given the relatively small size of black businesses compared to other businesses, blacks, particularly black youth, have expressed desire for more entrepreneurship courses in schools to enhance their entrepreneurial potential. As Kourilsky and Walstad (1998) advocated, entrepreneurship education improves the feasibility for entrepreneurship by increasing the knowledge of students, building confidence and promoting self-efficacy. It also develops perceived desirability for entrepreneurship by showing students that this activity is highly regarded and socially accepted by the community. Furthermore, Walstad and Kourilsky (1998) found that black minorities have greater interest in entrepreneurship and stronger desire to learn more about it as compared to whites.

Environmental entrepreneurs or "ecopreneurs" are entrepreneurs whose business focus is driven by both profit and concern for the environment. Ecopreneurship is an innovative, market-based approach that identifies and exploits opportunities to improve environmental quality (Anderson and Leal, 1997). Given the relatively high proportion of Americans who are concerned with the environment (three-fourths consider the environment a "high" priority and four out of five Americans consider themselves to be environmentalists), there are ample opportunities for ecopreneurs to make a difference in the

environmental sector (Baden and Noonan, 1996). Indeed, Samson (1994) suggested that educators in environmental and business fields should incorporate ecopreneurial-related lessons into their courses to help entrepreneurs identify a potential business opportunity when there is an environmental need or problem.

Notwithstanding the growing interest in entrepreneurship education for women, Blacks and ecopreneurs, there is a dearth of literature in these areas. It will be worthwhile for future studies to examine the issues pertaining to entrepreneurship education for women, Blacks, and ecopreneurs, which will help enhance the field's understanding of the contribution of entrepreneurship education for these niche groups.

## 8. CONCLUSION

There is a continuing interest in the field of entrepreneurship education and research in this area has grown rapidly over the years. Despite the growth in entrepreneurship education and training, there is little uniformity in the courses offered at all levels. Nonetheless, commentators in the field emphasized that nonuniformity in the courses offered is not necessarily a bad approach to program development as the key is to develop an effective mechanism that systematically documents the effects of those courses. While the literature distinguished between enterprise education and entrepreneurship education, there is a general agreement among researchers that the ultimate aim of both types of education is to encourage independent business creations. On the other hand, the literature indicates a consensus on the incompatibility of the curriculum of a business school with that of an entrepreneurship program. Although there has been much debate as to whether entrepreneurship can be taught, recent studies reported that both the scientific (i.e., functional areas of business) and creative aspects of entrepreneurship can be nurtured by entrepreneurship education. Empirical evidence indicates that entrepreneurship education is positively related to the creative facets of entrepreneurship such as the individual's adeptness in creating opportunities and ability to perceive the desirability and feasibility of a venture.

In terms of the teaching methods employed in entrepreneurship education and training, the learning approach has been touted as an effective path toward developing the entrepreneurial spirit. Entrepreneurship educators should stress in the curriculum, the importance of hands-on, active participation within a real life entrepreneurial environment where constructive feedback from an expert is provided. Evidence in the extant literature suggests that not only does entrepreneurship education positively influences individuals' propensity toward a more entrepreneurial stance, attitudes toward entrepreneurship education was also found to be positively related to new venture creation. Future research

should further examine the relationship between these factors. Additionally, given the dearth of literature in the areas of entrepreneurship education for women, black minorities, and ecopreneurs, future studies should explore these fields of research at greater length.

The new millennium is likely to involve greater environmental uncertainty and competition among businesses, resulting in a highly tumultuous economy that pressurizes government policy makers to increase the current stock of businesses. Job seekers too are not spared from this potential change. Flexibility and innovativeness will be critical survival skills in the highly competitive job market. Hence, it is vital that societies are encouraged to pursue entrepreneurial careers, and what is even more crucial is for universities and institutions of higher learning to provide courses and support to these potential entrepreneurs. Ultimately, this translates to the need for entrepreneurship education and training to be continuously monitored and evaluated to ensure that their objectives are met. It will be essential for evaluation studies to measure the pre- and post- test effects of entrepreneurship courses, and incorporate longitudinal research designs that use control groups to compare participants of entrepreneurship programs with individuals who did not have entrepreneurial educational experience.

## NOTES

<sup>1</sup> Please see Katz (2003) for a chronology of the growth of tertiary entrepreneurship education in the U.S.

<sup>2</sup> Entrepreneurship Awareness: General information programs on the reflection of entrepreneurship as a career option. Business Creation: Training in technical, human, and managerial skills to create a business. Small Business Development: Made-to-measure programs to answer the specific needs of owners/managers who cannot afford to pay specialist. Training of trainers: Programs to develop educators' skills in consultation, education, and follow-up of small businesses.

<sup>3</sup> There are currently more than 1,500 colleges and universities in the U.S. that offer some form of entrepreneurship training, more than 100 active university-based entrepreneurship centers, and more than 270 endowed professorships and chairs in entrepreneurship in the U.S. with an investment of nearly U.S. \$500 million. In addition, Robinson and Haynes (1991) reported that a new entrepreneurship center operates with a typical budget of U.S. \$2 million.

<sup>4</sup> The various business functions are finance, marketing, production, and human resource.

<sup>5</sup> Some of the examples of differences between the learning focus of business schools and entrepreneurship education are learning in the classroom vs. learning while and through doing; evaluation through written assessment vs. evaluation by judgment of people and events through direct feedback, success in learning measured by knowledge-based examination pass vs. success in learning by solving problems and learning from failure; critical judgment after analysis of large amounts of information vs. "gut feel" decision making with limited information; seeking the correct answer with time to do it vs. developing the most appropriate solution under pressure.

<sup>6</sup> Entrepreneurship consists of originating or starting a company while small business management consists of managing an existing company.

<sup>7</sup> “Unsupported Isolated Course” is the combination of a single course with low integration, while “Integration Program” is a combination of multiple courses that are well integrated both internally (with one another) and externally (with the curriculum).

<sup>8</sup> Other combinations proposed by the framework are “Unidisciplinary Approach” (single discipline, focusing on mature firms) and “Interdisciplinary Approach” (multiple disciplines, focusing on start-up firms).

<sup>9</sup> The “educators point of view” of program development is commonly found in academic institutions like universities, while student-based perspective is generally adopted when the priority lies in the psychosocial well-being of students.

<sup>10</sup> Stanford Technology Ventures Program (STVP) plays an active role at Stanford University in educating future scientists and engineers about high-technology entrepreneurship.

<sup>11</sup> It is important to note that teamwork teaching is more expensive, requires higher commitment from academicians, and it takes time for the partnership to succeed.

<sup>12</sup> These figures are based on a study by St. Louis University.

<sup>13</sup> These studies do not measure the pre- and post-test effects of entrepreneurship education, and they also lack control groups. Additionally, McMullan et al. (2001) reported that the commonly used subjective participant satisfaction measures are not correlated with objective measures of subsequent venture performance.

<sup>14</sup> Entrepreneurship graduates were 3 times more likely to start new businesses than general business graduates.

<sup>15</sup> McMullan and Gillin (1998) observed that 38% of the MBA Hybrid Entrepreneurship Program graduates at University of Calgary started new businesses beyond eight months of graduation as compared to 18% of the MBA program (no entrepreneurship courses) graduates at Australian University.

<sup>16</sup> Clouse (1990) highlighted that an introductory entrepreneurship course has a statistically significant impact on students’ simulated new venture decision behavior. About 75% of the participants developed positive decision behavior at the end of the semester as compared to at the inception of the course.

<sup>17</sup> Out of the 228 participants surveyed in the Quebec “Become an Entrepreneur” course, Garnier and Gasse (1990) found that 14% set up a business within 18 months after completing the entrepreneurship course and 51% actively pursued an entrepreneurial idea.

<sup>18</sup> Garnier et al. (1991) assessed that a year after attending a televised course in entrepreneurship, 12% of participants had launched a venture, 12% had developed the firm they owned before the course, and 44% had actively pursued their intentions to go into business.

<sup>19</sup> According to Price and Monroe (1992), entrepreneurs of the Fast Trac training program significantly increased their full-time employees and annual sales by 27% and 33%, respectively.

<sup>20</sup> Carrier (1999) reported that 80% of the entrepreneurs surveyed believed that training in export financing will assist them in their businesses, while more than half perceived that knowledge about the legal aspects of international trade will improve their export potential.

<sup>21</sup> For example, the Annual Conference organized by the Academy of Management (AOM)—Entrepreneurship Division attracts researchers and practitioners from both the U.S. and world-wide. As of May 2005, there are 1777 members in the Entrepreneurship Division, comprising of 1194 domestic (U.S.) and 583 international members. Out of the 1777 members, 1176 are academics, 19 are emeritus, 125 are executives and 457 are students.



<sup>22</sup> International Journal of Entrepreneurship Education (IJEE), an online journal dedicated to the publication of research papers and case studies on entrepreneurship and entrepreneurship education was introduced in 2002.

<sup>23</sup> For complete information on the Roundtable for Entrepreneurship Education, please visit the website <http://ree.stanford.edu>.

<sup>24</sup> REE USA conference in 2002 attracted over 75 members from 40 different universities.

<sup>25</sup> Please refer to note 14.

<sup>26</sup> Please refer to note 18.

<sup>27</sup> If further research indicates that positive attitudes toward entrepreneurship education do influence start-up activities, policy makers at both the government and academic levels should invest their efforts in changing the attitudes and mindsets of their target group and develop incentives for people to be genuinely interested and willing to commit their time to entrepreneurship education. The emphasis of entrepreneurship will shift from merely providing entrepreneurship courses and training to changing attitudes of would-be entrepreneurs. If business start-ups influence attitudes toward entrepreneurship education, business schools and colleges should target these groups of entrepreneurs who are interested in gathering knowledge and skills from entrepreneurship education and are willing to sacrifice their time, money and effort. While these people may have been successful in establishing businesses, the real pay-off concerns whether or not these businesses are successful in maintaining their position and growth in the existing market. Therefore, entrepreneurship education is still very much relevant to these entrepreneurs as the benefits of entrepreneurship education go beyond the start-up stage.

<sup>28</sup> Please visit <http://www.metrostate.edu/com/cwe/education.html> for further information on these programs.

## REFERENCES

- Acs, Z., Carlsson, B. and Karlsson, C. (1999). *Entrepreneurship, Small and Medium Enterprises and the Macro-economy*. Cambridge University Press, London.
- Anderson, T. L. and Leal, D. R. (1997). *Enviro-capitalists: Doing Good While Doing Well*. Rowman and Littlefield Publishers, Inc., Lanham.
- Baden, J. A. and Nooman, D. S. (1996). *The Evolving Environmentalism*. <Http://www.free-eco.org/free/pubs/DD.EE.html>.
- Béchar, J. P. and Toulouse, J. M. (1998). Validation of a Didactic Model for the Analysis of Training Objectives in Entrepreneurship. *Journal of Business Venturing* 13: 317–332.
- Birley, S. (1984). Education for Entrepreneurs: A U.K./USA Comparative Study. *Paper presented at the Small Business Institute Directors Association, Denver, Colorado*.
- Block, Z. and Stumpf, A. A. (1992). Entrepreneurship Education Research: Experience and Challenge. In *The State of the Art of Entrepreneurship*, D. L. Sexton and J. D. Kasarada (Eds), pp. 17–42. PWS-Kent Publishing Company, USA.
- Brockhaus, R. H. (1991). Entrepreneurship Education and Research Outside North America. *Entrepreneurship Theory and Practice* 15 (3): 77–84.
- Brown, L. E., Christy, R. L. and Banowetz, A. F. (1987). Perceptions of Success in Business Start-Up and the Impact of Entrepreneurial Education. *Frontiers of Entrepreneurship Research*, pp. 660–661. Babson College, Wellesley, MA.
- Brush, C. (1992). Research on Women Business Owners: Past Trends, New Perspective and Future Directions. *Entrepreneurship Theory and Practice* 16 (4): 5–30.

- Brush, C. G., Duhaime, I. M., Gartner, W. B., Stewart, A., Katz, J. A., Hitt, M. H., Alvarez, S. A., Meyer, G. D. and Venkataraman, S. (2003). Doctoral Education in the Field of Entrepreneurship. *Journal of Management* 29 (3): 309–331.
- Bruyat, C. and Julien, P. A. (2000). Defining the Field of Research in Entrepreneurship. *Journal of Business Venturing* 16 (2): 165–180.
- Carrier, C. (1999). The Training and Development Needs of Owner-Manager of Small Businesses with Export Potential. *Journal of Small Business Management* 37 (4): 30–41.
- Carsrud, A. L. (1991). Entrepreneurship and Enterprise Formation: A Brief Perspective on the Infrastructure in Europe. *Entrepreneurship Theory and Practice* 15 (3): 69–75.
- Center for Women's Business Research (2002). [Http://www.nfwb.org/index.asp](http://www.nfwb.org/index.asp).
- Chamard, J. (1989). Public Education: Its Effect on Entrepreneurial Characteristics. *Journal of Small Business and Entrepreneurship* 6 (2): 23–30.
- Charney, A. and Libecap, G. D. (2000). *Impact of Entrepreneurship Education: Insights*. Kauffman Research Series, Kauffman Center for Entrepreneurial Learning, USA.
- Clouse, Van. G. H. (1990). A Controlled Experiment Relating Entrepreneurial Education to Students' Start-Up Decisions. *Journal of Small Business Management* 28 (2): 45–54.
- Collins, O. F., Moore, D. G. and Umwalla, D. B. (1964). *The Enterprising Man*. MSU Business Studies, Bureau of Business and Economic Research, East Lansing, MI.
- Cox, L. W., Mueller, S. L. and Moss, S. E. (2002/2003). The Impact of Entrepreneurship Education on Entrepreneurial Self-Efficacy. *International Journal of Entrepreneurship Education* 1 (2).
- Curran, J. and Stanworth, J. (1989). Education and Training for Enterprise: Some problems of classification, evaluation, policy and research. *International Small Business Journal* 7 (2).
- Curran, J. and Storey, D. J. (2002). Small Business Policy in the United Kingdom: The Inheritance of the Small Business Service and Implications for its Future Effectiveness. *Environment and Planning C: Government and Policy* 20: 163–177.
- Dana, L. P. (1992). Entrepreneurial Education in Europe. *Journal of Education for Business* 68 (2): 74–78.
- Dana, L. P. (2001). The Education and Training of Entrepreneurs in Asia. *Education & Training* 43 (8/9).
- Donckels, R. (1991). Education and Entrepreneurship Experience from Secondary and University Education in Belgium. *Journal of Small Business and Enterprise* 9 (1): 35–42.
- Dyer, W. G. Jr. (1994). Toward a Theory of Entrepreneurial Careers. *Entrepreneurship Theory and Practice*, 19 (2): 7–21.
- Evans, D. S. and Leighton, L. (1990). Small Business Formation by Unemployed and Employed workers. *Small Business Economics* 2 (4): 319–320.
- Filion, L. J. (1994). Ten Steps to Entrepreneurial Teaching. *Journal of Small Business and Entrepreneurship* 5 (1): 27–31.
- Galloway, L. and Brown, W. (2002). Entrepreneurship Education at University: A Driver in the Creation of High Growth Firms? *Education and Training* 44 (8/9): 398–405.
- Garavan, T. N. and O'Connell, B. (1994). Entrepreneurship Education and Training Programs: A Review and Evaluation-Part 2. *Journal of European Industrial Training* 18 (11): 13–21.
- Garnier, B. and Gasse, Y. (1990). Training Entrepreneurs Through Newspapers. *Journal of Small Business Management* 28 (1): 70–73.
- Garnier, B., Gasse, Y. and Raynal, C. (1991). Evaluation of a Televised Course in Entrepreneurship. *Journal of Small Business and Entrepreneurship* 9 (1): 25–34.
- Gartner, W. B. and Vesper, K. H. (1994). Experiments in Entrepreneurship Education: Successes and Failures. *Journal of Business Venturing* 9: 179–187.
- Gasse, Y. (1985). A Strategy for the Promotion and Identification of Potential Entrepreneurs at the Secondary School Level. *Frontiers of Entrepreneurship Research*, pp. 538–559. Babson College, Wellesley, MA.

- Gibb, A. A. (1993). The Enterprise Culture and Education. Understanding Enterprise Education and its Links with Small Business Entrepreneurships and Wider Educational Goals. *International Small Business Management Journal* 11 (3).
- Gibb, A. A. (1996). Entrepreneurship and Small Business Management: Can We Afford to Neglect Them in the 21st Century Business School? *British Journal of Management* 7: 309–321.
- Gibb, Y. K. and Nelson, E. G. (1996). Personal Competences, Training and Assessment: A Challenge for Small Business Trainers. *Proceedings of the European Small Business Seminar*, pp. 97–107.
- Gnyawali, D. R. and Fogel, D. S. (1994). Environments for Entrepreneurship Development: Dimensions and Research Implications. *Entrepreneurship Theory and Practice* 19: 43–62.
- Gorman, G. and Hanlon, D. (1997). Some Research Perspectives on Entrepreneurship Education, Enterprise Education and Education for Small Business Management: A Ten-Year Literature Review. *International Small Business Journal* 15 (3): 56–78.
- Hansemark, O. C. (1998). The Effects of an Entrepreneurship Program on Need for Achievement and Locus of Control of Reinforcement. *International Journal of Entrepreneurial Behavior and Research* 4 (1): 28.
- Hansemark, O. C. (2003). Need for Achievement, Locus of Control, and the Prediction of Business Start-Ups: A Longitudinal Study. *Journal of Economic Psychology* 24 (3): 301.
- Hills, G. E. (1988). Variations in University Entrepreneurship Education: An Empirical Study of an Evolving Field. *Journal of Business Venturing* 3: 109–122.
- Hitt, M. A., Ireland, R. D., Camp, S. M. and Sexton, D. L. (2001). Guest Editors' Introduction to the Special Issue: Strategic Entrepreneurship: Entrepreneurial Strategies for Wealth Creation. *Strategic Management Journal* 22 (6/7): 479.
- Hult, G. T. M., Snow, C. C. and Kandemir, D. (2003). The Role of Entrepreneurship in Building Cultural Competitiveness in Different Organizational Types. *Journal of Management* 29 (3): 401.
- Hynes, B. (1996). Entrepreneurship Education and Training—Introducing Entrepreneurship into Non-Business Disciplines. *Journal of European Industrial Training* 20 (8): 10–17.
- Interman, H. (1992). International Management Development Network in Cooperation with the United Nations Development Program (UDP) and the International Labor Office (ILO). *Networking for Entrepreneurship Development*. International Labor Office, Geneva.
- Jack, S. L. and Anderson, A. R. (1998). *Entrepreneurship Education within the Condition of Entrepreneurology*. Proceedings of the Conference on Enterprise and Learning, Aberdeen.
- Jamieson, I. (1984). Schools and Enterprise. In *Education for Enterprise*, A. G. Watts and P. Moran (Eds), pp. 19–27. CRAC, Ballinger, Cambridge.
- Jennings, D. F. and Zeithamil, C. P. (1983). Locus of Control: A Review and Directions for Entrepreneurial Research. *Proceedings of the Academy of Management*, 417–421.
- Johannisson, B. (1991). University Training for Entrepreneurship: A Swedish Approach. *Entrepreneurship and Regional Development* 3 (1): 67–82.
- Kailer, N. (1990). Further Training in Small and Medium-sized Enterprises (Austria). *Journal of Small Business Management* 28 (1): 60–63.
- Kantor, J. (1988). Can Entrepreneurship Be Taught? A Canadian Experiment. *Journal of Small Business and Enterprise* 5 (4): 12–19.
- Kao, R. W. Y. (1994). From General Management to Entrepreneurship: The Business (B) School Challenge. *Journal of Small Business and Entrepreneurship* 11 (2): 4–10.
- Katz, J. A. (1991). Endowed Positions: Entrepreneurship and related fields. *Entrepreneurship Theory and Practice* 15 (3): 53–67.
- Katz, J. A. (2003). The Chronology and Intellectual Trajectory of American Entrepreneurship Education: 1876–1999. *Journal of Business Venturing* 18 (2): 283–300.

- Knight, R. M. (1987). Can Business Schools Produce Entrepreneurs? An Empirical Study. *Frontiers of Entrepreneurship Research*, pp. 603–604. Babson College, Wellesley, MA.
- Koch, L. T. (2002/2003). Theory and Practice of Entrepreneurship Education: A German View. *International Journal of Entrepreneurship Education* 1 (4).
- Kolb, D., Rabin, M. and McIntyre, J. M. (1974). *Organizational Philosophy: An Experiential Approach*. Prentice Hall, Englewood Cliffs, N.J.
- Kourilsky, M. L. and Walstad, W. B. (1998). Entrepreneurship and Female Youth: Knowledge, Attitudes, Gender differences and Educational Practices. *Journal of Business Venturing* 13 (1): 77–88.
- Kourilsky, M. L. and Walstad, W. B. (2002). The Early Environment and Schooling Experience of High-Tech Entrepreneurs: Insights for Entrepreneurship Education. *International Journal of Entrepreneurship Education* 1 (1): 1–20.
- Krueger, N. (1993). The Impact of Prior Entrepreneurial Exposure on Perceptions of New Venture Feasibility and Desirability. *Entrepreneurship Theory and Practice* 18 (1): 5–21.
- Krueger, N. and Brazeal, D. V. (1994). Entrepreneurial Potential and Potential Entrepreneurs. *Entrepreneurship Theory and Practice* 18 (3): 91–104.
- Krueger, N. F. and Dickson, P. R. (1994). How Believing in Ourselves Increases Risk Taking: Perceived Self-Efficacy and Opportunity Recognition. *Decision Science*, 25: 385–400.
- Laukkanen, M. (2000). Exploring Alternative Approaches in High-Level Entrepreneurship Education: Creating Micro-Mechanisms for Endogenous Regional Growth. *Entrepreneurship & Regional Development* 12: 25–47.
- Lee, L. and Wong, P. K. (2003). Attitudes toward Entrepreneurship Education and New Venture Creation. *Journal of Enterprising Culture* 11 (4): 339–357.
- Maillat, D. (1998). Innovative Milieux and New Generation of Regional Policies. *Entrepreneurship and Regional Development* 10: 1–16.
- Marlow, S. (1992). The Take-up of Business Growth Training Schemes by Ethnic Minority-owned Small Firms in Britain. *International Small Business Journal* 10 (4): 34–46.
- McCarthy, A., Morris, P. J. and Winn, J. (1997). A New Look at Undergraduate Entrepreneurship Education. *Proceedings of USABE Annual National Conference Entrepreneurship: The Engine of Global Economic Development, San Francisco, California*.
- McClelland, D. C. (1961). *The Achieving Society*. Van Nostrand, Princeton, N.J.
- McMullan, W. E. and Boberg, A. L. (1991). The Relative Effectiveness of Projects in Teaching Entrepreneurship. *Journal of Small Business and Entrepreneurship* 9 (9): 14–24.
- McMullan, W. E., Chrisman, J. J. and Vesper, K. (2001). Some Problems in Using Subjective Measures of Effectiveness to Evaluate Entrepreneurial Assistance Programs. *Entrepreneurship Theory and Practice* 26 (1): 37–54.
- McMullan, W. E. and Gillin, L. M. (1998). Industrial Viewpoint—Entrepreneurship Education. *Technovation* 18 (4): 275–286.
- McMullan, W. E. and Long, W. A. (1987). Entrepreneurship Education in The Nineties. *Journal of Business Venturing* 2: 261–275.
- McMullan, W. E., Long, W. A. and Graham, J. B. (1986). Assessing Economic Value Added by University-Based New-Venture Outreach Programs. *Journal of Business Venturing* 1: 225–240.
- McMullan, W. E., Long, W. A. and Wilson, A. (1985). MBA Concentration on Entrepreneurship. *Journal of Small Business and Entrepreneurship*. 3 (1): 18–22.
- Parker, S. C. (2005). Learning About the Unknown: How Fast Do Entrepreneurs Adjust Their Beliefs? *Forthcoming in Journal of Business Venturing*.
- Peterman, N. and Kennedy, J. (2003). Enterprise Education: Influencing Students' Perceptions of Entrepreneurship. *Entrepreneurship Theory and Practice* 28 (2): 129–144.

- Plaschka, G. R. and Welsch, H. P. (1990). Emerging Structures in Entrepreneurship Education: Curricular Designs and Strategies. *Entrepreneurship Theory and Practice* 14 (3): 55–70.
- Preshing, W. A. (1991). Education by Projects. *Journal of Small Business and Entrepreneurship* 9 (1): 55–59.
- Price, C. and Monroe, S. (1992). Educational Training for Women and Minority Entrepreneurs Positively Impacts Venture Growth. *Frontiers of Entrepreneurship Research*, pp. 216–230. Babson College, Wellesley, MA.
- Rae, D. and Carswell, M. (2001). Toward a Conceptual Understanding of Entrepreneurial Learning. *Journal of Small Business and Enterprise Development* 8 (2): 150–158.
- Robinson, P. and Haynes, M. (1991). Entrepreneurship Education in America's Major Universities. *Entrepreneurship Theory and Practice* 15 (3): 41–52.
- Ronstadt, R. (1985). Training Potential Entrepreneurs. In *Entrepreneurship: What it is and How to Teach It*, J. J. Kao and H. S. Stevenson (Eds). Harvard Business School, Boston.
- Rosa, P., Scott, M. G. and Klandt, H. (1996). *Educating Entrepreneurs in Modernizing Economies*. Avebury, Ipswich.
- Saeed, J. (1996). A Critical Evaluation of Australian Entrepreneurship Education and Training. *Proceedings of the Internationalizing Entrepreneurship Education and Training Conference*, Arnhem.
- Saks, N. T. and Gaglio, C. M. (2002). Can Opportunity Identification Be Taught? *Journal of Enterprising Culture* 10 (4): 313–347.
- Samson, K. J. (1994). Teaching Environmental Entrepreneurship. *Business* 16 (6): 32–34.
- Sandberg, W. R. and Gatewood, E. J. (1991). A Profile of Entrepreneurship Research Centers: Orientations, Interests, Activities, and Resources. *Entrepreneurship Theory and Practice* 15 (3): 11–24.
- Sexton, D. L. and Bowman, N. B. (1984). Entrepreneurship Education Suggestions for Increasing Effectiveness. *Journal of Small Business Management* 22 (2).
- Sexton, D. L. and Kasarda, J. D. (1992). *The State of the Art of Entrepreneurships*. P.W. Kent Publishing Co., Boston, MA.
- Sexton, D. L. and Upton, N. (1987). Evaluation of an Innovative Approach to Teaching Entrepreneurship. *Journal of Small Business Management* 25 (1): 35–43.
- Shepherd, D. A. and Douglas, E. J. (1997). Is Management Education Developing or Killing the Entrepreneurial Spirit? Proceedings of the 1997 USASBE Annual National Conference Entrepreneurship: *The Engine of Global Economic Development*, San Francisco, California.
- Singh, J. B. (1990). Entrepreneurship Education as a Catalyst of Development in the Third World. *Journal of Small Business and Entrepreneurship* 7 (4): 56–63.
- Solomon, G. T., Duffy, S. and Tarabishy, A. (2002). The State of Entrepreneurship Education in the United States: A Nationwide Survey and Analysis. *International Journal of Entrepreneurship Education* 1 (1): 1–22.
- Solomon, G. T. and Fernald, L. W. (1991). Trends in Small Business Management and Entrepreneurship Education in the United States. *Entrepreneurship Theory and Practice* 15 (3): 25–39.
- Stanworth, J. and Gray, C. (1992). Entrepreneurship and Education: Action-based: Research with Training Policy Implications in Britain. *International Small Business Journal* 10 (2): 11–23.
- Sternberg, R. and Wennekers, S. (2005). Determinants and Effects of New Business Creation: Using Global Entrepreneurship Monitor Data. *Small Business Economics* 24 (3): 193–204.
- Streeter, D. H., Jaquette Jr., J. P. and Hovis, K. (2002). *University-wide Entrepreneurship and Education: Alternative Models and Current Trends*. Working Paper No. 2002-02. Cornell University, USA.
- Stumpf, S. S., Dunbar, R. L. and Mullen, T. P. (1991). Simulations in Entrepreneurial Education: Oxymoron or Untapped Opportunity? *Frontiers of Entrepreneurship Research*, pp. 681–694. Babson College, Wellesley, MA.

- Vesper, K. H. (1971). Venture Initiation Courses in U.S Business Schools. *Academy of Management Journal* 6: 14–19.
- Vesper, K. H. (1985). New Developments in Entrepreneurship Education. *Frontiers of Entrepreneurship Research*, pp. 489–497. Babson College, Wellesley, MA.
- Walstad, W. B. and Kourilsky, M. L. (1998). Entrepreneurial Attitudes and Knowledge of Black Youth. *Entrepreneurship in Theory and Practice* 23 (2): 5–18.
- Wong, P. K., Ho, Y. P. and Autio, E. (2005). Entrepreneurship, Innovation and Economic Growth: Evidence from GEM data. *Small Business Economics* 24 (3): 335–350.
- Zeithamil, C. P. and Rice, G. H. (1987). Entrepreneurship/Small Business Education in American Universities. *Journal of Small Business Management* 25 (1): 44–50.

HERBERT J. SCHUETZE<sup>a</sup> and HEATHER ANTECOL<sup>b</sup>

<sup>a</sup>*Department of Economics, University of Victoria, Victoria, BC, Canada*  
*E-mail: hschuetz@uvic.ca*

<sup>b</sup>*Department of Economics, Claremont McKenna College, Claremont, California, USA*  
*E-mail: heather.antecol@claremontmckenna.edu*

## 5. Immigration, Entrepreneurship and the Venture Start-up Process

### 1. INTRODUCTION<sup>1</sup>

Immigrants represent an important source of human capital that is central to the development of an entrepreneurial base for many countries. Immigrants are indeed an entrepreneurial lot—with self-employment rates among many immigrant groups that are significantly higher than those of native workers. For example, Fairlie (Chapter 15 of this handbook) finds that all racial groups in the U.S., with the exception of Latino immigrants, have higher self-employment rates than natives. Similar results have been identified elsewhere in the literature for the U.S. as well as other countries (Clark and Drinkwater, 1998 (U.K.), Borjas, 1986; Fairlie and Meyer, 1996; Fairlie and Woodruff, 2004 (U.S.) and Schuetze, 2005 (Canada)). The potential to invigorate entrepreneurship through immigration has not gone unnoticed by policy makers. Several countries have implemented programs designed to attract immigrant entrepreneurs. Immigration policies in countries such as Australia, Canada and Germany include special visas and entry requirements that facilitate immigration by would-be entrepreneurs.

Perhaps because of the importance of self-employment among immigrants, a number of research studies have attempted to identify the reasons for the high incidence of this labor market activity among immigrants compared to natives. A number of researchers have suggested that cultural factors related to one's country of origin may play a role in determining immigrants' higher propensities toward self-employment. Light (1984) found significant heterogeneity in self-employment rates across country of origin and attributed

it to diversity in traditions of commerce. Light and Rosenstein (1995) suggest that different ethnic groups are endowed with diverse “supply characteristics” that may provide a special advantage in entrepreneurship; such as the ability to cook Chinese food among Chinese immigrants. Others have suggested that attitudes toward entrepreneurship may be related to one’s religion (Carrol and Mosakowski, 1987; Rafiq, 1992; Clark and Drinkwater, 2000). However, empirical support for the hypothesis that self-employment rates among immigrant groups are correlated with home country self-employment rates has been mixed. While Yeungert (1995) found that immigrants from countries with high self-employment rates are more likely to become self-employed in the U.S., Fairlie and Meyer (1996) found no correlation.

One of the dominant explanations for the high rates of immigrant self-employment put forth early in the literature is blocked mobility. It has been argued that ethnic minorities faced with discrimination from employers in the wage and salary sector turn to self-employment as a way to advance in such segmented labor markets<sup>2</sup> (Light, 1972; Sowell, 1981; Moore, 1983, Min, 1984; Phizacklea, 1988; Metcalf, Modood and Virdee, 1996; Mata and Pendakur, 1998; Li, 1998; Wong and Ng, 2002). Examples include Min (1984) and Wong and Ng (2002) who found that among Korean entrepreneurs in Atlanta and Chinese self-employed in Vancouver, respectively, disadvantage in the nonethnic market was a major reason for choosing self-employment.

Another leading hypothesis suggests that the presence of ethnic concentrations or “enclaves” in the host country create opportunities for potential immigrant entrepreneurs. It is argued that ethnic enclaves provide its members with greater access to capital through the pooling of investment resources (Light, 1972), a supply of local labor (Light and Bonacich, 1988) and consumers with tastes for goods that ethnic entrepreneurs are better positioned to provide (Aldrich et al., 1985). However, evidence for such an enclave effect has also been mixed. As Parker (2004, pp. 121–122) points out, while some studies find that the presence and size of ethnic enclaves positively impacts the probability of self-employment among its members (Le, 2000; Flota and Mora, 2001; Lofstrom, 2002), many do not (Borjas and Bronars, 1989; Yuengert, 1995; Razin and Langlois, 1996; Clark and Drinkwater, 1998, 2000, 2002). Even within studies there is not consensus. For example, Borjas (1986) and Boyd (1990) find support for an enclave effect among some ethnic groups but not others.

Despite the relatively large literature explaining differences in the *levels* of self-employment, however, few studies (only three to our knowledge: Borjas, 1986; Lofstrom, 2002; and Schuetze, 2005) have examined the *start-up process* and the factors that influence the early years of the business life cycle among self-employed immigrants.<sup>3</sup> This gap in the literature is important to fill because characterizing the start-up process may provide important clues to how and why



immigrant firms are formed and develop with years since migration in the host country. In addition, all previous studies focus on immigrant self-employment outcomes in a single country, which limits opportunities to examine the role that immigration policy plays in immigrant self-employment.

In this chapter we establish the key features of the venture start-up process among immigrants and attempt to identify factors that influence their decision to start a business. Primarily due to data limitations our working definition of entrepreneurs throughout this chapter includes all individuals who are “self-employed” and we use the terms interchangeably. Because our focus in this chapter is the path into self-employment among immigrants, the primary outcome of interest is whether or not immigrants choose self-employment (the propensity toward self-employment). However, in an attempt to determine the “quality” of immigrant entrepreneurial outcomes we also examine their weekly earnings. To help identify common traits across countries in the start-up process and the role that immigration policy and country-specific institutions/market factors play we analyze self-employment outcomes in three countries—Australia, Canada and the United States. The economies in these countries are quite similar in many respects, as has been documented in the literature, but (as we discuss in the next section) differ in important ways with respect to labor market institutions, broad immigration policy and policy with respect to immigrant entrepreneurship. All of this creates a good “natural experiment” with which to compare the immigrant experience.<sup>4</sup>

The availability of high-quality census microdata from the three countries allows us to examine a comparable and detailed analysis of the start-up process. Because cross-section studies of immigrant outcomes confound secular changes in cohort outcomes with changes in the start-up patterns of self-employment within cohorts,<sup>5</sup> we employ an empirical approach, similar to the relatively large number of studies examining immigrant outcomes in the wage sector, to examine the self-employment outcomes of immigrants. In particular, pairs of data files are used to perform an empirical decomposition that allows identification of start-up patterns (changes in self-employment incidence within an arrival cohort) for comparison across immigrant cohorts of different vintages. This method also allows for identification of the performance of self-employed immigrants relative to natives.

The remainder of this chapter is as follows. Section 2 highlights the immigration policies and other institutional/structural market differences relevant to immigrant self-employment outcomes across the three countries over the period examined. Section 3 outlines the empirical strategy used to identify the key elements of immigrants’ self-employment experience. Section 4 describes the data and our primary estimation results (those focusing on business start-up), while Section 5 examines the “quality” of self-employment outcomes by examining earnings. Finally, Section 6 concludes the chapter.

## 2. IMMIGRATION POLICY, STRUCTURAL MARKET CHARACTERISTICS AND SELF-EMPLOYMENT

Australia, Canada and the United States have a long tradition of immigration that, by international standards, has resulted in large immigrant population shares. The similarities and differences in the immigration experiences of these three countries are well known and have been exploited by a number of researchers to analyze the impacts of immigration policy and institutional/market factors on immigrant outcomes.<sup>6</sup> This section provides a comparison of the immigration policies, institutions, and structural market characteristics across these three countries.<sup>7</sup> Particular attention is given to those differences that are likely relevant to self-employment outcomes among immigrants and to the period in which cohort assimilation profiles can be determined from the data utilized (roughly 1956–1990). For a more exhaustive history of these countries' immigration policies see, for example, Antecol, Cobb-Clark and Trejo (2003) for Australia; Green (1995), Green and Green (1992; 1995) for Canada; and for a comparison between Canadian and U.S. policy, see Borjas (1993).

To begin, we describe similarities and differences in immigration policy and their likely impacts on self-employment outcomes. Until the 1960s, entry into the three countries was based primarily on national origin. In the United States the composition of visas distributed was set to match the national origin of the foreign-born population of the 1920 U.S. Census. Canadian policy gave preference to immigrants from Britain, northwest Europe and the United States, while Australia's favored British immigration (Antecol, Cobb-Clark, Trejo, 2003). In 1962 Canada moved away from national origin as the right to sponsor family members for immigration was extended to nontraditional source countries. The U.S. also largely abandoned national origin (Briggs, 1984); a move that was later followed by Australia in 1973, which ended its "White Australia" policy. The policies that replaced national origin differed significantly across the three countries. The introduction of "points systems" in 1967 in Canada and in the late 1970's in Australia placed more weight on skilled migration. Under these points systems, nonsponsored immigrants enter under the "skilled category" and are evaluated and obtain "points" based on observable skills which are considered important to one's success in the labor market. Those obtaining enough points are permitted entry. In contrast, the U.S. adopted an immigration policy that placed greater emphasis on family reunification.

The shift away from national origin led to a significant change in the source regions of immigrants to all three countries (see Green, 1995 for Canada; Borjas, 1993 for the U.S.; and Miller, 1999 for Australia). The composition of immigrants was increasingly comprised of immigrants from Asia, for example,

as opposed to the more traditional regions of the United Kingdom and Western Europe. In addition, while there is considerable debate in the literature regarding the precise reason for the changes in skill composition,<sup>8</sup> it has been shown that a shift in the relative skill levels of immigrants arriving across the three countries occurred subsequent to these policy changes. As is shown by Borjas (1993), who compares Canada and the United States and Chiswick (1987) who includes Australia in his comparison of immigrant outcomes, these shifts resulted in average education levels among immigrants arriving to Canada and Australia<sup>9</sup> which were higher than those among immigrants to the United States following the changes in policy in the three countries.<sup>10</sup>

While these shifts in composition likely had differential impacts on the self-employment outcomes of immigrants to these countries, the precise nature of these impacts is unclear. Given that immigrant cohorts were increasingly comprised of ethnic minorities in all three countries, the blocked mobility hypothesis suggests that self-employment propensities among cohorts arriving after the move away from national origin would have increased. However, as noted above, the probability of self-employment may also be affected by the level of self-employment in the source countries of immigrants and the presence of ethnic enclaves in the receiving country. Moreover, there are substantial differences in the ethnic composition of immigrants across the three countries (see Antecol, Cobb-Clark and Trejo 2003a, 2003b), thus it is unclear what impact the move away from national origin likely had on relative self-employment outcomes in these three countries. In addition, previous research largely suggests that education or skill level is positively correlated with the probability of self-employment in the overall population.<sup>11</sup> Thus, increases in the skill level relative to the U.S. among immigrants arriving in Canada and Australia may have led to increases in self-employment in these two countries.

The introduction of the skilled class of workers to Canada and Australia likely also had a more direct impact on self-employment among immigrants. In 1976 in Australia and 1978 in Canada, a second category of skilled worker was added with the creation of a “business skills” class in both countries. The business immigrant programs introduced in Australia and Canada were similar in many respects but differed in a number of important ways. During the period covered by this study both programs contained three sub-categories of business class immigrants; the “business owner,” “senior executive” and “investor”<sup>12</sup> categories were introduced in Australia and the “entrepreneur,” “self-employment,” and (in 1986) “investor” categories in Canada. The primary differences between the business skills classes and the “standard” skilled worker class are the criteria used to assess workers’ skills and the requirement to run a business subsequent to entry. Like the Canadian program, the Australian business skills program placed greater weight on previous experience (business experience) and the availability of investment funds<sup>13</sup> than the other skilled

immigrant entry category. In addition, both countries' business skill programs required immigrants, by threat of visa cancellation, to enter into business within the first three years after arrival.<sup>14</sup>

However, much like the other skilled immigrant categories,<sup>15</sup> the criteria used to determine permanent residency under the business skills program was much more stringent in Australia than in Canada. Also, unlike Canada's program, the criteria for entry through the business class in Australia were more difficult to satisfy than those of the other skilled classes. In general, in order to be eligible for permanent residence in Australia immigrants had to fulfill a more demanding set of requirements with respect to age, qualifications, experience and language ability than those intending to migrate to Canada. In particular, with respect to business immigrants, the Australian program required that applicants had owned or operated a business for at least two of the four years prior to application. This was not required of business immigrants entering Canada.

This more stringent assessment of potential immigrant entrepreneurs is likely to result in fewer immigrants choosing self-employment upon entry to Australia relative to Canada (and perhaps even the United States, which has no similar entry program), but increased longevity among permanent immigrants who do start businesses. On the other hand, the requirement to operate a business within the first few years after arrival may have resulted in higher rates of self-employment immediately after entry relative to the U.S. Thus, the impacts of the Australian business immigrant program on self-employment early after entry relative to Canada were likely negative while the impacts relative to U.S. entry self-employment rates are ambiguous.

Indeed, Figure 5-1, which illustrates the percent of total immigration that is comprised of business immigration in Australia and Canada up until 1990, bears this out to a certain extent. While the percentages of total immigration comprised of business immigrants entering Australia and Canada annually had increased substantially by the late 1980's, these rates were lower in Australia in every year than those in Canada. Between 1980 and 1986, Canada received an annual inflow of nearly 6000 business class immigrants; accounting for an average of 6% of overall immigration. By the late 1980s the number of business class immigrants had increased to over 15,000 annually or 9% of overall immigration. In comparison, in 1982, six years after the introduction of the program and the first year for which we have data, just over 1000 immigrants, or 1%, of total immigrants entered under the business immigrant program in Australia. This number grew to 10,000 immigrants or 8.25% of the total number of Australian immigrants by 1989. These numbers likely reflect the more selective nature of the business immigrant program in Australia relative to Canada.

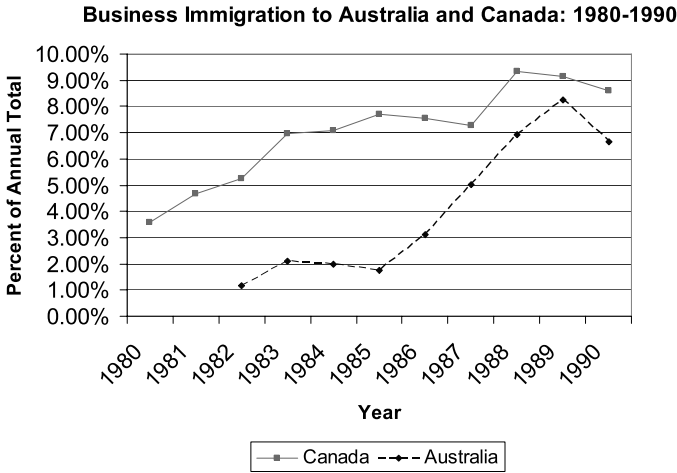


FIGURE 5-1 Business immigration to Australia and Canada. Sources: Canadian data—Citizenship and Immigration Canada (various years) “Citizenship and Immigration Statistics,” Cat. No. MP22-1; Minister of Public Works and Government Services Canada. Australian data—Department of Immigration and Multicultural and Indigenous Affairs (various years), “Population Flows: Immigration Aspects,” Australian Bureau of Statistics.

Aside from differences in immigration policy, structural and institutional differences across the three countries are also likely to influence the type of immigrants attracted to each destination and, therefore, their self-employment outcomes. More generous social programs (including national health insurance, unemployment insurance and welfare systems) in Australia and Canada are likely to attract workers at the lower end of the skill distribution. While the literature on the impacts of taxation is mixed (see Parker, 2004, and Schuetze, 2004, for reviews), this type of selection is likely reinforced by the fact that income tax rates in the United States are structured in such a way as to allow successful entrepreneurs to reap more of their benefits relative to those in Canada and Australia. The one cross-country study by Schuetze (2000), which examines differences in taxation between Canada and the U.S., supports this hypothesis. In addition, the access to large markets in the United States may attract previously successful entrepreneurs from other countries. The availability and generosity of loan guarantee schemes, direct loans to small businesses and other government programs, such as those that encourage self-employment among the unemployed in Canada and the U.S., may result in differences in immigrant self-employment outcomes across the three countries. Our comparative analysis will help to shed light on the relative importance of the selection of immigrants based on structural/institutional factors and that of immigration policy.

### 3. EMPIRICAL FRAMEWORK

We adopt the regression framework developed by Borjas (1986), as set out in Schuetze (2005), for estimating the effects of duration in destination country (number of years since migration) and year of arrival cohorts on self-employment propensities. Focusing on the group of employed immigrants, the following cross-sectional self-employment probability equation can be estimated by probit for each country separately:

$$P_t = \Phi \left( X_t' \beta_t + \sum_{k=1}^K \delta_{kt} \right), \quad (1)$$

where  $P_t$  is the probability of self-employment in year  $t$ ,  $X_t$  is a vector of observable characteristics related to the self-employment decision,  $\Phi$  is the normal cumulative distribution function,  $k$  indexes a series of five-year arrival cohorts identified by the earliest year of arrival among those in the cohort (e.g., the 1971–75 cohort is labeled 1971), and the  $\delta_{k,t}$  are cohort-year specific intercepts.

One could estimate changes in the self-employment propensity of immigrants through time using the coefficient estimates of the  $\delta_{k,t}$  from (1) and looking across cohorts of different vintages. However, such estimates are unreliable if time varying cohort specific fixed effects impact the two sectors of employment differently. Borjas (1995) and Baker and Benjamin (1994) find that more recent cohorts of immigrants to the U.S. and Canada, respectively, have poorer earnings outcomes in the wage sector than earlier cohorts. Thus, the use of a single cross-section is unlikely to be appropriate. Instead, estimates which are free of this potential fixed effect bias can be obtained using quasi-panel methods in which “synthetic cohorts” of immigrants are followed through time. Therefore, (1) is estimated at two points in time (1981 and 1991 in Australia and Canada, and 1980 and 1990 in the U.S.).

With this estimation strategy a decomposition of the cross-section change in the probability of self-employment can be stated as follows. Consider the predicted probability of self-employment for a cohort group  $k$  in 1991 evaluated at the average values of immigrant characteristics in that year ( $\bar{X}$ ). This probability is given by:

$$\hat{P}_{k,91} = \Phi(\bar{X} \hat{\beta}_{91} + \hat{\delta}_{k,91}). \quad (2)$$

The predicted probability of the cohort in 1991 that arrived ten years later than  $k$ , again evaluated at  $\bar{X}$ , is:

$$\hat{P}_{k+10,91} = \Phi(\bar{X} \hat{\beta}_{91} + \hat{\delta}_{k+10,91}). \quad (3)$$

Given the definitions in (2) and (3) the cross section change in the self-employment propensity over ten years for 1991 is equal to  $\hat{P}_{k,91} - \hat{P}_{k+10,91}$ . Following Borjas (1985) this change can be decomposed into two components as follows:

$$\hat{P}_{k,91} - \hat{P}_{k+10,91} = (\hat{P}_{k,91} - \hat{P}_{k,81}) + (\hat{P}_{k,81} - \hat{P}_{k+10,91}), \quad (4)$$

where  $\hat{P}_{k,81}$  is the predicted probability of self-employment for cohort  $k$  in 1981 evaluated at  $\bar{X}$  (i.e., at the average values of immigrant characteristics in 1991). More specifically, it is the following prediction:

$$\hat{P}_{k,81} = \Phi(\bar{X}\hat{\beta}_{81} + \hat{\delta}_{k,81}). \quad (5)$$

It is important to note that cohort  $k$  in 1981 has the same number of years since migration as cohort  $k + 10$  in 1991.

Thus, the first term on the right-hand side of (4) gives the change in the predicted percent of immigrants in cohort  $k$  experiencing self-employment over the ten year period.<sup>16</sup> In other words, these estimates provide a measure of net self-employment start-up for each of our immigrant cohort groups. Because we observe immigrant cohorts of all vintages (from just after arrival onward) these estimates paint a picture of the entire self-employment “start-up process” for immigrants.<sup>17</sup> The second term in (4) gives the difference in the probability of self-employment between two cohorts with the same number of years since migration. This difference provides an estimate of the impact of cohort fixed effects on the propensity to choose self-employment. It will be negative if more recent cohorts are more likely to choose self-employment (e.g., if wage employment outcomes are worse relative to self-employment for more recent cohorts).

An additional concern arises when there are unobserved time effects (other than those arising from years in host country) that change over the ten year period. In this case the estimates in (4) are biased. A common solution to this problem in the literature is to normalize the changes in immigrant outcomes to some base group. Our base group is comprised of native workers. Thus, we also provide estimates of the within and across cohort changes in self-employment probabilities among immigrants that are net of changes in these predicted probabilities among native workers<sup>18</sup> with similar characteristics over the ten-year period (see Schuetze, 2005, for a more detailed account).

#### 4. ESTIMATION AND RESULTS

The data used in the analysis are drawn from the 1981 and 1991 (1% samples) Australian Censuses, the 1981 (2% sample) and 1991 (3% sample)

Canadian Census Public Use Microdata Files,<sup>19</sup> and the U.S. Census, 5% public use A samples for 1980 and 1990. These data files are chosen because they are the most recent census pairs available covering the same period and for which sufficient comparable information is available to carry out the analysis.<sup>20</sup> The samples are restricted to males who are employed in the survey week (the week prior to the survey), who are not in the armed forces<sup>21</sup> and not in school at the time of the survey. To control for aging within cohorts across the ten-year time frames, the samples are restricted to individuals aged 18–54 in 1980/81 and to those aged 28–64 in 1990/91.<sup>22</sup> Because of the prevalence of self-employment in the agricultural sector among nonimmigrants the samples are further restricted to individuals employed in nonagricultural industries.<sup>23</sup> Due to the large sample sizes of the U.S. census data, 40% random samples of nonimmigrants are taken. Weights are applied throughout the calculations to the U.S. samples that account for the unbalanced samples taken and the fact that the 1990 Census is a nonrandom sample of the population. We also exclude individuals with missing values for any of the variables used in the analysis. Finally, the data files are pooled across pairs of census files in each country.

The primary outcome variable of interest in the analysis is an indicator of self-employment activity. The self-employment indicator used is based on the class of worker variables in Canada and the United States and on occupational/labor force status in Australia. In both Canada and the United States, the definition of self-employed includes individuals who indicate that they work for themselves in incorporated or unincorporated businesses and those in professional practices. In Australia, the definition of self-employed includes individuals who indicate they are conducting their own business irrespective of whether they employ others. In all three countries, the self-employed definitions exclude unpaid family workers. Table 5-1 describes how self-employment varies with nativity and immigrant arrival cohort in the three countries.<sup>24</sup> Here, and throughout the chapter, the intervals listed for immigrant arrival cohorts are those defined in the Australian and Canadian data; the slightly different immigrant cohorts defined in the U.S. data are as follows: pre-1960, 1960–64, 1965–69, 1970–74, 1975–79, 1980–84 and 1985–90.<sup>25</sup> The 1991 Australian census does not distinguish 1960s arrivals from earlier immigrants, and therefore “pre-1971” is the most precise arrival cohort that can be defined consistently across censuses for Australian immigrants. For Canada and the United States, however, immigrants arriving during these years are disaggregated into “1966–70,” “1961–65” and “pre-1961” cohorts.

Table 5-1 shows that in the United States, immigrants as a group have average self-employment rates that are only slightly above those of native workers (with immigrant-native self-employment differentials of approximately 1 percentage point), whereas in Canada and Australia immigrant men tend to have considerably higher self-employment rates than their native-born coun-



TABLE 5-1 *Self-employment rates of men*

|            | Australia                   |                             | Canada                      |                              | U.S.                         |                              |
|------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|------------------------------|------------------------------|
|            | 1981                        | 1991                        | 1981                        | 1991                         | 1980                         | 1990                         |
| Natives    | 0.136<br>(0.343)<br>[17389] | 0.179<br>(0.384)<br>[21068] | 0.099<br>(0.299)<br>[62379] | 0.117<br>(0.321)<br>[104923] | 0.102<br>(0.302)<br>[727852] | 0.110<br>(0.312)<br>[788906] |
| Immigrants | 0.153<br>(0.360)<br>[6344]  | 0.204<br>(0.403)<br>[7591]  | 0.145<br>(0.352)<br>[17016] | 0.163<br>(0.369)<br>[26325]  | 0.114<br>(0.318)<br>[130510] | 0.117<br>(0.321)<br>[221569] |
| Cohort     |                             |                             |                             |                              |                              |                              |
| Pre-1961   |                             |                             | 0.177<br>(0.309)<br>[68976] | 0.192<br>(0.394)<br>[4459]   | 0.158<br>(0.364)<br>[32093]  | 0.161<br>(0.367)<br>[22925]  |
| 61-65      |                             |                             | 0.147<br>(0.354)<br>[1960]  | 0.195<br>(0.397)<br>[2150]   | 0.139<br>(0.346)<br>[17348]  | 0.148<br>(0.355)<br>[17561]  |
| 66-70      | 0.171<br>(0.377)<br>[4686]  | 0.240<br>(0.427)<br>[3430]  | 0.140<br>(0.347)<br>[3539]  | 0.176<br>(0.380)<br>[4792]   | 0.123<br>(0.328)<br>[22825]  | 0.142<br>(0.349)<br>[23476]  |
| 71-75      | 0.108<br>(0.310)<br>[837]   | 0.182<br>(0.386)<br>[1053]  | 0.114<br>(0.318)<br>[3100]  | 0.162<br>(0.369)<br>[5018]   | 0.100<br>(0.301)<br>[27351]  | 0.143<br>(0.350)<br>[30027]  |
| 76-80      | 0.094<br>(0.292)<br>[821]   | 0.247<br>(0.432)<br>[777]   | 0.087<br>(0.282)<br>[1820]  | 0.158<br>(0.365)<br>[3483]   | 0.061<br>(0.239)<br>[30893]  | 0.129<br>(0.335)<br>[36242]  |
| 81-85      |                             | 0.174<br>(0.379)<br>[876]   |                             | 0.148<br>(0.355)<br>[2675]   |                              | 0.101<br>(0.301)<br>[45988]  |
| 86-91      |                             | 0.129<br>(0.336)<br>[1455]  |                             | 0.108<br>(0.310)<br>[3748]   |                              | 0.059<br>(0.236)<br>[45350]  |

## Notes:

- Samples in all years are restricted to men aged 18-54 (20-54 in Australia).
- Standard deviations are given in parentheses.
- Number of observations is given in brackets.

terparts (immigrant-native self-employment differentials of approximately 4 and 2 percentage points in both years, respectively). In Australia and Canada, male self-employment rates rose for both natives and immigrants between 1981 and 1991, although the increases were larger in Australia (increases of 4.3 and 5.1 percentage points for natives and immigrants, respectively) than in Canada (where increases for both groups were 1.8 percentage points). In the U.S., male rates of self-employment remained relatively stable over the 1980-90 period. The fact that the increases across nativity were quite similar within

countries, despite differences across the three countries, might suggest that country-specific factors play a role in determining self-employment outcomes among immigrants. Finally, examining the raw self-employment rates across cohorts shows that more recent arrival cohorts in all years/countries have lower rates of self-employment than earlier arrivals and natives. However, within 5–10 years in Canada, and slightly longer (10–15 years) in the U.S. and Australia, immigrants have completely closed the immigrant/native self-employment gap, irrespective of the survey year. This might suggest that some immigrants require time in the destination country before starting up businesses and/or that more recent arrivals are less likely to become entrepreneurs. Distinguishing between these two explanations is only possible by identifying within cohort changes in self-employment separately from across cohort changes and controlling for differences in characteristics across groups.

As indicated above, such decompositions are given by (4) net of the outcomes of our base group. In order to derive these, we estimate probit models of the probability of self-employment, separately, for each country for immigrants and natives jointly using the pooled data files. In identifying demographic and economic characteristics that influence the relative returns in the two sectors for inclusion in the analysis we are guided by previous research on the determinants of self-employment<sup>26</sup> and immigrant outcomes in the wage and salary sector.<sup>27</sup> We include controls for age, level of education, marital status and source country of immigrants (see the Appendix for variable definitions).

To make the decomposition more tractable we place some constraints on the coefficients. In particular, we allow the impacts of the demographic and economic variables to differ across immigrants and natives but restrict them to be the same within these groups across the years examined; such that;  $\beta_{81} = \beta_{91} = \beta$ . With these assumptions the decompositions for each country can be derived using the pooled data files through the estimation of a single probit equation and interacting the appropriate coefficients with indicator variables for survey year and immigration status to match the constraints imposed. To overcome the classic problem of distinguishing between cohort, age and period effects we impose the common identifying restriction that the period effect is the same for immigrants and natives.<sup>28</sup>

Table 5-2 highlights some of the trends in predicted outcomes derived from the coefficient estimates.<sup>29</sup> The first row provides the predicted probability of self-employment for each year and country for a native with the average characteristics of immigrants in 1990/91 in each country. Consistent with overall trends in self-employment in the three countries, these results show secular growth in self-employment among these representative workers in Australia and Canada and little or no change in the United States. Much like the raw trends, self-employment grew most rapidly in Australia even after controlling

TABLE 5-2 Summary of selected trends in predicted values

|                       | Australia |         | Canada  |         | U.S.    |         |
|-----------------------|-----------|---------|---------|---------|---------|---------|
|                       | 1981      | 1991    | 1981    | 1991    | 1980    | 1990    |
| Predicted probability | 0.182     | 0.235   | 0.136   | 0.148   | 0.122   | 0.127   |
| Natives <sup>a</sup>  | (0.004)   | (0.003) | (0.002) | (0.001) | (0.000) | (0.000) |
| GAP <sup>b</sup>      | -0.052    | -0.071  | -0.022  | -0.029  | -0.038  | -0.045  |
|                       | (0.014)   | (0.013) | (0.009) | (0.006) | (0.002) | (0.002) |

## Notes:

Standard errors derived from bootstrapping are given in parentheses.

<sup>a</sup>The predicted probability of self-employment for a native with similar characteristics as the average immigrant in 1990/91.

<sup>b</sup>The predicted probability “gap” in the self-employment rates between the most recent arrival cohort and a similar native.

for individual characteristics. Row 2 of the Table 5-2 gives the predicted self-employment rate gap between a representative immigrant from the most recent arrival cohorts in each year/country and a similar native. This gap is negative and statistically significant in all cases, which suggests that immigrants to all three countries initially enter with self-employment rates that are lower than those among similar native workers. Consistent with our expectations on the impacts of a more selective business immigrant program in Australia, the gaps in self-employment rates upon entry are greatest among immigrants to Australia. These gaps are lowest among immigrants to Canada, perhaps as a result of the requirement to own and operate a business early after entry to Canada (within the first three years) under the business skills program, although the Canada-U.S. differences are not large. Finally, while the gap grew between 1980/81 and 1990/91 in all three countries, the gap grew most substantially (from 5 percentage points to 7 percentage points) in Australia while in Canada and the U.S. it grew more modestly, from 2.2 (3.8) to 2.9 (4.5) percentage points in Canada (the U.S.). To determine whether or not these gaps persist over time in the destination country we turn next to the decompositions.

Table 5-3 presents decompositions of the cross-section self-employment propensity profiles into estimates of the “within” cohort increases in self-employment net of cohort effects, our measure of the self-employment start-up process and estimates of the effects of changes in cohort propensities “across” cohorts with similar years since migration. The first column under each of the country headings gives the cross-section prediction of self-employment growth while the second and third columns provide the decompositions of these without adjusting for secular changes in self-employment over the period. The last two columns account for secular changes in self-employment in each of the countries using natives as the base group. The cross-section estimates show little growth in business start-up among immigrants beyond the first 10–15

TABLE 5-3 *Decomposition of changes in the probability of self-employment*

|       | Australia  |         |          |         |               |         | Canada     |         |          |         |               |         | U.S.       |         |          |         |               |         |         |         |
|-------|------------|---------|----------|---------|---------------|---------|------------|---------|----------|---------|---------------|---------|------------|---------|----------|---------|---------------|---------|---------|---------|
|       | Unadjusted |         | Adjusted |         | Cross-section |         | Unadjusted |         | Adjusted |         | Cross-section |         | Unadjusted |         | Adjusted |         | Cross-section |         |         |         |
|       | Within     | Across  | Within   | Across  | Within        | Across  | Within     | Across  | Within   | Across  | Within        | Across  | Within     | Across  | Within   | Across  | Within        | Across  |         |         |
| 51-61 |            |         |          |         |               |         |            |         |          |         |               |         |            |         |          |         |               |         |         |         |
|       | 0.061      | -0.091  | 0.007    | -0.037  | -0.003        | 0.023   | 0.009      | -0.005  | -0.002   | 0.007   | -0.010        | 0.000   | 0.034      | -0.014  | 0.022    | -0.002  | 0.003         | -0.018  | -0.002  | -0.013  |
|       | (0.010)    | (0.019) | (0.011)  | (0.020) | (0.008)       | (0.009) | (0.007)    | (0.007) | (0.007)  | (0.007) | (0.003)       | (0.003) | (0.011)    | (0.011) | (0.011)  | (0.011) | (0.004)       | (0.003) | (0.004) | (0.003) |
| 66-76 | 0.061      | -0.091  | 0.007    | -0.037  | -0.003        | 0.023   | 0.009      | -0.005  | -0.002   | 0.007   | -0.010        | 0.000   | 0.034      | -0.014  | 0.022    | -0.002  | 0.003         | -0.018  | -0.002  | -0.013  |
|       | (0.019)    | (0.010) | (0.011)  | (0.020) | (0.008)       | (0.009) | (0.007)    | (0.007) | (0.007)  | (0.007) | (0.003)       | (0.003) | (0.011)    | (0.011) | (0.011)  | (0.011) | (0.004)       | (0.003) | (0.004) | (0.003) |
| 71-81 | 0.020      | -0.072  | 0.038    | -0.018  | 0.012         | 0.039   | 0.039      | -0.026  | 0.027    | -0.015  | 0.024         | 0.036   | 0.039      | -0.026  | 0.027    | -0.015  | 0.024         | 0.036   | -0.011  | 0.031   |
|       | (0.022)    | (0.021) | (0.021)  | (0.021) | (0.010)       | (0.009) | (0.009)    | (0.010) | (0.009)  | (0.010) | (0.003)       | (0.003) | (0.009)    | (0.010) | (0.009)  | (0.010) | (0.003)       | (0.003) | (0.003) | (0.003) |
| 76-86 | 0.110      | -0.035  | 0.091    | 0.019   | 0.057         | 0.061   | 0.061      | -0.004  | 0.049    | 0.008   | 0.069         | 0.067   | 0.061      | -0.004  | 0.049    | 0.008   | 0.069         | 0.067   | 0.002   | 0.062   |
|       | (0.021)    | (0.018) | (0.023)  | (0.019) | (0.029)       | (0.011) | (0.011)    | (0.010) | (0.011)  | (0.010) | (0.002)       | (0.003) | (0.011)    | (0.010) | (0.011)  | (0.010) | (0.002)       | (0.003) | (0.002) | (0.003) |

Notes:

For a description of the decomposition see text. Standard errors derived from bootstrapping are given in parentheses.

years after arrival in all three countries. Most of the entries are statistically insignificant with the exception of those for the most recent arrivals. However, the “unadjusted” results of the decomposition show that, once across cohort effects are taken into account, immigrants in all three countries experience significant within cohort increases in business start-up that extend beyond the first 10–15 years in the destination country. The across cohort estimates, which are negative and generally statistically significant, suggest that secular changes in the composition of immigrant cohorts have led to increases in self-employment propensities among immigrants with similar years in all three countries. This confirms that the cross-section self-employment growth estimates are biased.

Looking across countries (still focusing on the unadjusted results), an interesting pattern of self-employment growth is observed through time in the destination country. In all countries, the within cohort increases in self-employment propensities are higher among the most recent arrival cohorts than among those with more potential labor market experience in the destination country. Thus, it appears that the business start-up process is accelerated in the first 10–15 years after arrival. In addition, the rate of growth within cohorts is much higher across all vintages of immigrants in Australia than in Canada or the U.S. This is interesting in light of the fact that the self-employment rate gap between immigrants and natives upon entry to Australia is much larger than those in Canada and the U.S. (see Table 5-2). It also appears that the across cohort increases in self-employment were higher in Australia. These patterns may also reflect overall trends in employment compensation in the three countries that have made self-employment more attractive, a dominant feature of the Australian experience. Before deriving any conclusions, however, the analysis must control for general trends in self-employment outcomes.

The “adjusted” entries in each of the country panels account for these secular trends. The overall effect of normalizing the results is to dampen both the within and across cohort effects (albeit only very slightly in the U.S.). In fact, a number of the entries in the “adjusted” column become small in magnitude and statistically insignificant. This suggests that some of the growth in self-employment within and across cohorts is explained by secular increases in the probability of self-employment within the three countries. Despite this, some of the entries remain significant after normalizing and interesting patterns of self-employment “assimilation” continue to emerge across the countries.

First, there is some evidence of increases in the propensity toward self-employment across cohorts that are consistent with the timing of changes in immigration policy that were implemented between the 1960’s and the 1970’s in all three countries. The move away from national origin occurred early in the United States, the effects of which were most noticeable by the early 1970’s in terms of source country composition. The statistically significant

across cohort entries for the “61–71” and the “66–76” cohorts are consistent with the timing of these changes and with the shift in policy resulting in higher rates of self-employment among immigrants. As noted above, the move away from national origin occurred somewhat later in Australia and Canada. However, despite similar changes in policy with respect to national origin and the introduction of points systems, there is only weak evidence of across cohort increases in self-employment around the time of these changes in Canada and Australia. Point estimates for the “66–76” and “71–81” across cohort changes indicate somewhat sizable increases in the self-employment propensity for both countries; however, they are not statistically significant at conventional levels.

Second, it appears that a number of immigrants require time in wage employment before starting self-employment ventures. Rates of self-employment within cohorts net of secular trends increase with years since migration among immigrants to all three countries. Like the patterns observed in the unadjusted results, much of this “assimilation” occurs in the first 10–15 years after arrival. The estimated within cohort increases among the most recent arrivals in each of the three countries are larger in magnitude and more likely to be statistically significant than those among earlier arrival cohorts. For example, the most recent arrivals to Australia in the 1981 census (the 1976–1980 arrivals) experienced an increase of 9 percentage points over and above that predicted for similar natives in their first 10–15 years in that country between 1981 and 1991. This compares to a less than 4 percentage point net increase over this period among those who arrived five years earlier (between 1971 and 1975) and zero net growth for those who arrived ten years earlier.

In addition, the magnitudes of the increases in self-employment propensities experienced among immigrant cohorts in their first 10–15 years in each country are substantial. Continuing with Australia as an example, the gap between the “76” cohort and natives upon entry in 1981 was just over five percentage points (see Table 5-2). As noted above, this cohort experienced an estimated net increase in the self-employment rate of 9 percentage points and implies that by 1991 the rate of self-employment for this cohort was nearly 4 percentage points higher than a similar native. Similarly, the projected self-employment rate differentials between immigrants to Canada and the U.S. and similar natives were 2.8 and 2.4 percentage points based on the gap from Table 5-2 and within growth for the “76” cohorts in Table 5-3. In all cases immigrant self-employment rates caught up to and overtook those of similar natives in the first 10–15 years after arrival. Interestingly, this “overtaking” occurred despite significant differences in the size of the entry gaps across the countries. Indeed, it appears that the amount of “assimilation” that occurred after entry, to a certain extent, coincided with the size of the entry gaps. The entry gap and the net increases in self-employment subsequent to entry were highest among immigrants to Australia, followed by those to the United States,

and then those to Canada. However, consistent with our expectation of longer survival in self-employment due to the relatively more selective Australian skilled worker categories (both business and standard skilled classes), the amount of “overtaking” was less among immigrants to Canada and the United States than among Australian immigrants.

The fact that immigrant self-employment rates rapidly overtook those of natives despite differences in the general level of self-employment across the three countries suggests that, regardless of differences in the institutional or country specific factors that influence rates of self-employment, immigrants adapt quite rapidly. That immigrant self-employment rates eventually exceed those of similarly skilled natives may suggest that workers who choose to immigrate are more “entrepreneurial” than non-migrants. On the other hand (as discussed earlier in the introduction to this chapter), immigrants may choose self-employment not because they possess strong business skills but because of blocked mobility in the wage sector and, thus, self-employment represents employment of last resort. We attempt to sort between these two possibilities in the next section.

## 5. EARNINGS OUTCOMES

The results in Section 4 highlight the differences in the numbers of immigrants who choose self-employment across the three countries but give little indication as to whether or not these are good business ventures. In this section we attempt to shed light on the relative “quality” of the self-employment experiences of immigrants to Australia, Canada and the United States by examining the earnings outcomes of the self-employed. Once again using these three countries as a “laboratory” by analyzing this measure of quality we hope to further our understanding of the impacts of immigrant policy and other institutional/market characteristics on immigrant self-employment outcomes.

The approach taken to examine earnings is similar to that outlined above in Section 3 to examine self-employment propensities, except for the following important differences. First, with respect to the sample, we continue to focus on males in the same age categories as above and who are not in agricultural industries or enrolled in school at the time of the survey. However, we now restrict attention to individuals who are self-employed and with a reasonable attachment to the labor force (worked 14 or more weeks in the previous year).<sup>30</sup> Because the class of worker variables in the Canadian and U.S. census files refer to the type of employment during the survey week while the earnings data pertain to the previous year, we alter our definition of “self-employed” in the earnings analysis.<sup>31</sup> In particular, the data in these two countries contain information on the amount and source of income in the year

prior to the survey which we use to impute whether the individual was self-employed or not. For the vast majority of individuals in our samples, assigning self-employment status on the basis of this information is straightforward: most respondents had only one source of labor market income (wage and salary earnings or self-employment income). For those with multiple sources of earnings, we simply assigned individuals earning a substantial fraction<sup>32</sup> of their previous year's income from running a business to the "self-employed" category. Because income from the operation of an incorporated business is included in the wage and salary earnings of these individuals in these two surveys, unlike above, the definition of "self-employed" used in this section includes only those who operated unincorporated businesses. Given that most businesses in the early stages of development are likely to be unincorporated, this difference is likely not an issue.

The measure on which we concentrate our attention is the log of weekly earnings.<sup>33</sup> This measure was chosen, in part, because the hours worked variable in the Australian data is reported as a categorical variable; making hourly earnings infeasible to calculate. In addition, the Canadian data on hours pertains to the survey week rather than the year prior to the survey. To examine the sensitivity of our results to possible variation in hours worked we redid the analysis restricting the sample to full-time workers. These results, which are available upon request from the authors, were substantively similar to those presented in the chapter. Finally, because the earnings measure is continuous, unlike above, the estimation equations in this section are linear and estimated using multiple regression analysis.<sup>34</sup>

Table 5-4 describes how log weekly earnings among the self-employed vary with nativity and immigrant arrival cohort in the three countries. To enable the reader to draw comparisons across years, within each country, we restrict attention to individuals aged 18–54 (20–54 in Australia) in all years and inflate the 1980/1981 earnings to 1990/91 values using the CPI from each of the countries. However, comparisons of the levels of earnings across countries are not meaningful as we did not adjust the figures for the rates of exchange between the various currencies. On average, it appears that self-employed immigrants earn approximately the same amount of income as natives in all three countries. In addition, immigrants and natives in all three countries experienced little growth in real log earnings over the period. One difference across the three countries of note is that the pattern across cohorts in average earnings appears to differ in the United States from those observed in Australia and Canada. There are generally no significant differences between the average earnings of more recent and earlier arrival cohorts in Australia and Canada while raw average earnings increase with years in the United States.

The decomposition of the log weekly earnings of self-employed immigrants into within and across cohort changes presented in Table 5-5 allow U.S.



TABLE 5-4 Average log weekly earnings of self-employed men

|            | Australia                  |                            | Canada                     |                            | U.S.                        |                             |
|------------|----------------------------|----------------------------|----------------------------|----------------------------|-----------------------------|-----------------------------|
|            | 1981                       | 1991                       | 1981                       | 1991                       | 1980                        | 1990                        |
| Natives    | 6.155<br>(0.578)<br>[2147] | 6.125<br>(0.749)<br>[3298] | 6.408<br>(0.984)<br>[4323] | 6.325<br>(1.060)<br>[8133] | 6.283<br>(0.944)<br>[55798] | 6.168<br>(1.086)<br>[64674] |
| Immigrants | 6.146<br>(0.547)<br>[863]  | 6.097<br>(0.740)<br>[1338] | 6.417<br>(0.933)<br>[1392] | 6.312<br>(1.070)<br>[2361] | 6.361<br>(0.958)<br>[10347] | 6.196<br>(1.087)<br>[17831] |
| Cohort     |                            |                            |                            |                            |                             |                             |
| Pre-1961   |                            |                            | 6.472<br>(0.899)<br>[649]  | 6.375<br>(0.945)<br>[440]  | 6.462<br>(0.894)<br>[3446]  | 6.402<br>(1.034)<br>[2466]  |
| 61-65      |                            |                            | 6.301<br>(1.068)<br>[175]  | 6.348<br>(1.035)<br>[213]  | 6.440<br>(0.970)<br>[1630]  | 6.364<br>(1.020)<br>[1747]  |
| 66-70      | 6.138<br>(0.529)<br>[708]  | 6.075<br>(0.778)<br>[712]  | 6.388<br>(0.905)<br>[288]  | 6.415<br>(1.013)<br>[440]  | 6.351<br>(0.965)<br>[1922]  | 6.281<br>(1.142)<br>[2178]  |
| 71-75      | 6.177<br>(0.504)<br>[84]   | 6.126<br>(0.680)<br>[166]  | 6.422<br>(0.977)<br>[191]  | 6.317<br>(1.174)<br>[473]  | 6.294<br>(0.964)<br>[1951]  | 6.300<br>(1.076)<br>[2875]  |
| 76-80      | 6.184<br>(0.739)<br>[71]   | 6.081<br>(0.772)<br>[174]  | 6.332<br>(0.885)<br>[89]   | 6.217<br>(1.054)<br>[325]  | 5.124<br>(1.027)<br>[1398]  | 6.206<br>(1.112)<br>[3219]  |
| 81-85      |                            | 6.161<br>(0.604)<br>[126]  |                            | 6.285<br>(1.050)<br>[258]  |                             | 6.018<br>(1.034)<br>[3341]  |
| 86-91      |                            | 6.134<br>(0.693)<br>[160]  |                            | 6.094<br>(1.232)<br>[212]  |                             | 5.836<br>(1.080)<br>[2005]  |

## Notes:

- Samples in all years are restricted to men aged 18-54 (20-54 in Australia).
- Earnings values inflated to each countries 1990/91 level using CPI.
- Standard deviations are given in parentheses.
- Number of observations is given in brackets.

to control for possible changes in the “quality” of immigrants across cohorts, differences in labor market characteristics and secular trends in earnings. The second and third columns in each of the country panels, which give the decomposition results without adjusting for secular trends among natives, indicate significant increases both within and across cohorts in all three countries. Most of these increases, as indicated by the “adjusted” results, are attributable to growth in nominal earnings and other secular trends. Controlling for these secular trends (columns 4 and 5), we find that an interesting pattern emerges

TABLE 5-5 *Earnings regressions—assimilation, cohort and period effects*

|       | Australia     |        |            |        |          |        | Canada        |        |            |        |          |        | U.S.          |        |            |        |          |         |         |         |         |         |         |         |         |         |
|-------|---------------|--------|------------|--------|----------|--------|---------------|--------|------------|--------|----------|--------|---------------|--------|------------|--------|----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|       | Cross-section |        | Unadjusted |        | Adjusted |        | Cross-section |        | Unadjusted |        | Adjusted |        | Cross-section |        | Unadjusted |        | Adjusted |         |         |         |         |         |         |         |         |         |
|       | Within        | Across | Within     | Across | Within   | Across | Within        | Across | Within     | Across | Within   | Across | Within        | Across | Within     | Across | Within   | Across  |         |         |         |         |         |         |         |         |
| 51-61 |               |        |            |        |          |        | -0.024        | 0.441  | -0.422     | 0.006  | 0.012    | 0.057  | 0.476         | -0.506 | 0.058      | -0.001 | (0.750)  | (0.000) | (0.00)  | (0.917) | (0.697) | (0.086) | (0.000) | (0.000) | (0.022) | (0.971) |
| 61-71 |               |        |            |        |          |        | 0.122         | 0.667  | -0.545     | 0.232  | -0.110   | 0.059  | 0.486         | -0.484 | 0.068      | -0.009 | (0.594)  | (0.000) | (0.000) | (0.023) | (0.230) | (0.053) | (0.000) | (0.000) | (0.041) | (0.775) |
| 66-76 | -0.010        | 0.671  | -0.681     | -0.008 | -0.002   | 0.165  | 0.165         | 0.595  | -0.430     | 0.160  | 0.005    | 0.066  | 0.488         | -0.456 | 0.070      | -0.003 | (0.873)  | (0.000) | (0.000) | (0.834) | (0.974) | (0.219) | (0.000) | (0.000) | (0.000) | (0.905) |
| 71-81 | 0.001         | 0.669  | -0.668     | -0.010 | 0.011    | 0.024  | 0.024         | 0.503  | -0.479     | 0.068  | -0.043   | 0.198  | 0.567         | -0.417 | 0.149      | 0.048  | (0.994)  | (0.000) | (0.000) | (0.900) | (0.890) | (0.232) | (0.000) | (0.000) | (0.000) | (0.087) |
| 76-86 | 0.072         | 0.674  | -0.602     | -0.005 | 0.077    | 0.176  | 0.176         | 0.432  | -0.256     | -0.003 | 0.179    | 0.246  | 0.605         | -0.389 | 0.186      | 0.059  | (0.392)  | (0.000) | (0.000) | (0.961) | (0.474) | (0.098) | (0.000) | (0.000) | (0.000) | (0.134) |

Note: *P*-values are reported in parentheses.

across the countries. Male self-employed immigrants to Australia enter with average log weekly earnings that are about the same as a similar native (the estimated “gap” in 1981 is negative 3% but statistically insignificant) and earnings do not appear to increase relative to natives with time in the country. Looking across cohorts with the same number of years in Australia, there is no indication of changes in earnings outcomes; in other words, no indication of a change in quality across cohorts.

Compared to Australia, the outcomes among self-employed immigrants to Canada appear to be worse. The entry earnings gap between self-employed immigrants and natives (estimated at negative 21% and significant in 1981) suggests that an immigrant arriving in the late 1970’s entered with earnings below those of similarly skilled natives. In addition, while earnings among earlier cohorts increased relative to natives between 1981 and 1991, those of immigrants arriving to Canada in the late 1970’s did not. Neither the “71” or “76” arrival cohorts experienced significant within growth net of secular trends over the ten year period while the “61” and “66” arrival cohorts did. Thus, it appears that either the self-employed immigrant earnings outcomes of more recent arrivals are poorer relative to earlier arrivals or that it takes several years in Canada for the earnings of self-employed immigrants to catch up to those of natives. Finally, evidence of a further deterioration in earnings outcomes among Canadian immigrants is reflected in the earnings gap which fell to almost negative 40% in 1991.

In contrast to the earnings outcomes of self-employed immigrants to Canada, immigrants to the U.S. performed quite well relative to natives in that country. While male self-employed immigrants to the U.S. entered with earnings below those of similar natives (the gap is estimated at negative 18% in 1980), their earnings appear to catch up to those of natives in the first number of years in the country. Earnings among the “76” cohort grew by 18.6 percentage points relative to natives in the first 10–15 years in the United States. In fact, cohorts of all vintages to the United States in our sample experienced significant increases in earnings net of secular trends. This suggests that average earnings among self-employed immigrants to this country eventually surpass those of natives.

The variation in the earnings outcomes among self-employed immigrants suggests that there is a great deal of heterogeneity in the quality of the self-employment experiences across the three countries. That self-employed immigrants did not experience earnings outcomes that exceeded those of natives in Australia and Canada suggests that the higher rates of self-employment experienced among immigrants to all three countries is not likely due to greater business skills among immigrants to all countries. Instead, it appears that immigrants, much like natives, enter self-employment for varied reasons, which depend, in part, on country-specific factors. Lastly, the differences across

countries suggest that immigration policy and other country specific factors likely influence the quality of immigrant self-employment outcomes (discussed in detail in the subsequent section).

## 6. CONCLUSIONS

Given the recent emphasis placed on immigrant entrepreneurship by government policy makers around the world it is important to understand the potential for immigration policy in attracting successful entrepreneurs. A central issue underlying the likely success of immigration policy in achieving these goals is the ability of such policies to overcome other institutional and market forces that make some countries more attractive to entrepreneurs than others. We characterize the business start-up process for immigrant men and look for clues to the likely impacts of immigration policies and other institutional/market frameworks on immigrant self-employment outcomes by examining the self-employment experiences of immigrants to three countries: Australia, Canada and the United States. These three countries are similar in many respects but differ substantially with respect to immigration policies, other institutions, and market characteristics, which impact self-employment outcomes.

First, with respect to the characteristics of the business start-up process among immigrant men we find that a number of interesting conclusions can be drawn from the observed self-employment patterns. The results in all three countries suggest that the process of starting a business for many immigrants involves a transition from wage employment to self-employment. Immigrants to all three countries had self-employment rates below those of similar natives at the time of entry to the destination country. However, in all years/countries we find positive and statistically significant growth in the self-employment propensities of newly arriving immigrants over and above that of similar natives. This may be because, relative to wage employment, self-employment typically requires bigger financial investment, the development of contacts and greater country-specific knowledge. Given that all of these typically take time to acquire, perhaps it is not surprising that a period of integration is required.

In addition, despite very different rates of self-employment across the three countries, we find that rates of self-employment among immigrants catch up to and overtake those of similar natives within 10–20 years after arrival. This suggests that, regardless of differences in the institutional or country-specific factors that influence rates of self-employment, immigrants adapt to these conditions relatively quickly. This result also casts doubt on explanations for the observed higher rates of self-employment among immigrants, which suggest that immigrants do not assimilate but simply adopt their host-country self-employment propensities. While immigrant self-employment

rates eventually exceed those of similarly skilled natives, an examination of the earnings outcomes in self-employment across the three countries reveals that these higher self-employment rates may not be associated with greater entrepreneurial skill levels among immigrants arriving to all countries. We find a great deal of heterogeneity across the three countries in the earnings outcomes of immigrants relative to natives. For example, while the relatively “good” earnings outcomes among immigrants to the United States are consistent with immigrants possessing higher levels of business skill, the poor earnings outcomes among immigrants to Canada are not.

Second, with respect to the relative impacts of policy and other country specific factors on self-employment outcomes, we find evidence that, while immigration policy may affect self-employment outcomes at the margin, the most substantial determinants are likely other institutional/market structure forces that attract entrepreneurs. To see how we come to this conclusion, consider the differences in immigration policy and other market structures across the three countries. Australia and Canada are perhaps most alike among the three in terms of immigration policy, tax policy, size of local markets and other market factors. Unlike the United States, both of these countries have immigration policies that screen immigrants to a certain extent based on skill characteristics and formal business skill programs. There is, however, one primary difference between Australia and Canada’s immigration policies. Australia’s skills requirements for entry through both the business and other skill categories are relatively more stringent than those of Canada. Thus, a comparison between the self-employment outcomes of immigrants across Australia and Canada allows us to isolate the impact of this policy difference. Comparing Australia and Canada to the United States, on the other hand, provides information not only on the impacts of immigration policy (in particular, the presence or absence of a “points” system) but also the effects of differences in institutional/market characteristics. The United States differs from Australia and Canada in terms of the size of the local market, tax policy and other institutional factors. In particular, while Canada and Australia have more generous social programs, the U.S. has more favorable tax provisions for entrepreneurs under the income tax system and access to larger markets. These differences likely favor the United States relative to Canada and Australia as a destination of choice for the most skilled entrepreneurs.

As noted above, looking across countries we do indeed find evidence that suggests that immigration policy has an impact on self-employment outcomes. These impacts, perhaps not surprisingly, were most evident in the Australian results. For example, Australia’s relatively rigorous “points” requirements for entry appear to have had the expected effects both in terms of self-employment business start-up and earnings outcomes. Consistent with our expectations, relative to those to Canada and the United States, immigrants to

Australia entered with self-employment rates that were further below those of similarly skilled natives than those in the other two countries but experienced relatively high rates of self-employment with time in the destination country. Also consistent with a more selective points system, relative to those in Canada the earnings outcomes among male self-employed immigrants to Australia were more favorable. On the other hand, comparing immigrant self-employment rates in Canada to those in the U.S., we find little evidence that Canadian immigration policy has had a significant impact.

As a final point, our examination of the earnings outcomes among immigrants to the United States and comparison to those to Australia and Canada leads us to conclude that self-selection among immigrant entrepreneurs based on other market factors, such as market size and tax policy, are likely more important than immigration policy. Self-employed immigrants to the United States out-performed immigrants to Canada and Australia in terms of earnings outcomes relative to natives. These differences in the relative earnings outcomes among male immigrants between the United States and the other two countries were substantial, despite the fact that immigrants to the United States were not formally screened based on skills.

## APPENDIX

TABLE 5-6 *Variable definitions*

|                    | Australia                | Canada                    | U.S.                         |
|--------------------|--------------------------|---------------------------|------------------------------|
| age                | Age                      | Age                       | Age                          |
| agesq              | Age squared              | Age squared               | Age squared                  |
| Years of education |                          |                           |                              |
| ed1                | = 1 if less than 9 years | = 1 if less than grd. 5   | = 1 if grade 8 or less       |
| ed2                | = 1 if 10–13 years       | = 1 if grade 5–8          | = 1 if less than high-school |
| ed3                | = 1 if some college      | = 1 if grade 9–13         | = 1 if highschool grad       |
| ed4                | = 1 if BA+               | = 1 if sec. school grad   | = 1 if some college/bach.    |
| ed5                | NA                       | = 1 if trade certificate  | = 1 if masters/Ph.D.         |
| ed6                | NA                       | = 1 if non-univ: no trade | NA                           |
| ed7                | NA                       | = 1 if non-univ: trade    | NA                           |
| ed8                | NA                       | = 1 if non-univ: other    | NA                           |
| ed9                | NA                       | = 1 if univ: no cert      | NA                           |
| ed10               | NA                       | = 1 if univ: cert < bach  | NA                           |
| ed11               | NA                       | = 1 if bachelors+         | NA                           |
| Marital status     |                          |                           |                              |
| mstat1             | = 1 if single            | = 1 if divorced           | = 1 if married               |
| mstat2             | = 1 if married           | = 1 if married            | = 1 if widowed               |

TABLE 5-5 (Continued)

|                 | Australia                                   | Canada                           | U.S.                         |
|-----------------|---|----------------------------------|------------------------------|
| mstat3          | = 1 if<br>sep./div./widowed                 | = 1 if separated                 | = 1 if divorced              |
| mstat4          | NA  | = 1 if never married             | = 1 if separated             |
| mstat5          | NA  | = 1 if widowed                   | = 1 if never married         |
| Y91             | = 1 if 1991                                 | = 1 if 1991                      | = 1 if 1990                  |
| Place of birth  |   |                                  |                              |
| pob2            | = 1 if North America and<br>USA             | = 1 if USA                       | = 1 if Africa                |
| pob3            | = 1 if Germany                              | = 1 if UK                        | = 1 if Canada                |
| pob4            | = 1 if Netherlands                          | = 1 if Germany                   | = 1 if other North Am.       |
| pob5            | = 1 if UK and Ireland                       | = 1 if Italy                     | = 1 if Mexico                |
| pob6            | = 1 if Yugoslavia                           | = 1 if Portugal                  | = 1 if S. and Central<br>Am. |
| pob7            | = 1 if Italy                                | = 1 if Poland                    | = 1 if East Asia             |
| pob8            | = 1 if Southern Europe                      | = 1 if USSR                      | = 1 if South Asia            |
| pob9            | = 1 if Poland                               | = 1 if other Europe              | = 1 if Middle East           |
| pob10           | = 1 if USSR & Baltic<br>States              | = 1 if Asia                      | = 1 if other Asia            |
| pob11           | = 1 if other Europe                         | = 1 if Africa                    | = 1 if Western Europe        |
| pob12           | = 1 if Southeast Asia                       | = 1 if Central./South<br>America | = 1 if Southern Europe       |
| pob13           | = 1 if South Asia & India                   | = 1 if other                     | = 1 if Eastern Europe        |
| pob14           | = 1 if Mid East & North<br>Africa           | NA                               | = 1 if Northern Europe       |
| pob15           | = 1 if Africa                               | NA                               | = 1 if former USSR           |
| pob16           | = 1 if South/Central<br>America & Caribbean | NA                               | = 1 if Oceania               |
| pob17           | = 1 if Oceania,<br>Antarctica, NZ           | NA                               | = 1 if other                 |
| immig           | = 1 if immigrant                            | = 1 if immigrant                 | = 1 if immigrant             |
| Year of arrival |   |                                  |                              |
| D56             | NA  | = 1 if before 1961               | = 1 if before 1960           |
| D61             | NA  | = 1 if 1961–1965                 | = 1 if 1960–1964             |
| D66             | = 1 if before 1971*                         | = 1 if 1966–1970                 | = 1 if 1965–1969             |
| D71             | = 1 if 1971–1975*                           | = 1 if 1971–1975                 | = 1 if 1970–1974             |
| D76             | = 1 if 1976–1980*                           | = 1 if 1976–1980                 | = 1 if 1975–1979             |
| D81             | = 1 if 1981–1985                            | = 1 if 1981–1985                 | = 1 if 1980–1984             |
| D86             | = 1 if 1986–1991                            | = 1 if 1986–1991                 | = 1 if 1985–1990             |

\*In 1981 these categories for Australia are before 1972, 1972–1976, and 1977–1981, respectively.

## NOTES

<sup>1</sup> Funding for this research was provided by SSHRC (grant # 410-2003-1823). We are grateful to Hui Feng for competent research assistance. We are solely responsible for any omissions or errors.

<sup>2</sup> See Light and Rosenstein (1995) and Parker (2004) for a more thorough discussion of the literature.

<sup>3</sup> Borjas (1986) and Lofstrom (2002) examine data from the U.S. However, because these studies are limited to a single country over a limited period of time they do not provide much insight into the roles of immigration policy and other institutional/market factors in the start-up process—a topic of focus in this chapter. Schuetze (2005), which most closely resembles the current study, focuses on self-employment outcomes of immigrants through time in both Canada and the United States.

<sup>4</sup> A number of studies have made use of this setting. See, for example, Chiswick (1987), Antecol, Cobb-Clark, and Trejo (2003a, 2003b), and Antecol, Kuhn and Trejo (2003).

<sup>5</sup> See, for example, Borjas (1985) or LaLonde and Topel (1992).

<sup>6</sup> Examples include Chiswick (1987), Duleep and Regets (1992), Borjas (1993), and Antecol, Cobb-Clark and Trejo (2003).

<sup>7</sup> Much of this section is based on previous work by Antecol, Cobb-Clark and Trejo (2003).

<sup>8</sup> The debate, (see Duleep and Sanders, 1992; Borjas, 1993; and Antecol, Cobb-Clark and Trejo, 1993), centers around the issue of whether or not it was the introduction of the points systems in Canada and Australia which led to the shift in observable skills (such as, education levels) of immigrants. The fact that the shifts in skill occurred, however, is of central interest to the current chapter, not the reason for the shift per se.

<sup>9</sup> Perhaps because of the more stringent evaluation of skilled immigrants to Australia, Chiswick (1987) finds the average levels of education to be highest among Australian immigrants.

<sup>10</sup> Antecol, Cobb-Clark, and Trejo (2003a, 2003b) find similar results using the 1990/91 censuses for Australia, Canada and the U.S. However, once immigrants from Central/South America are excluded, the skill levels of immigrants are similar across the three countries.

<sup>11</sup> See Parker (2004, p. 73) for a review of this literature.

<sup>12</sup> Other categories are available to those who are sponsored by an Australian State/Territory. For more information see [www.immi.gov.au](http://www.immi.gov.au).

<sup>13</sup> A minimum amount of investment capital is required for those entering under the investor and entrepreneur categories.

<sup>14</sup> This is not the case for “investors” in either country.

<sup>15</sup> Lester and Richardson (2004) provide a good comparison of the two countries’ immigration policies.

<sup>16</sup> For this estimate to be unbiased it must be assumed that cohort specific fixed effects are equal across time. This may not be true in this setting if, for example, the composition of the cohort changes through the remigration of immigrants based on skills.

<sup>17</sup> By “start-up” we mean not necessarily that the firm itself is newly created (we do not observe this) but that the immigrant owner is new to a given venture in the destination country.

<sup>18</sup> However, as in Schuetze (2005), the results here are similar when the base group is comprised of previous immigrants. These results are available from the authors upon request.

<sup>19</sup> Because detailed information on year of arrival for immigrants in regions determined to have too few observations to protect confidentiality are unavailable, data drawn from the Canadian census files is restricted to Quebec, Ontario and the Western Provinces.

<sup>20</sup> In particular, the year of arrival information in the 2001 Australian Census is not detailed enough to analyze using the empirical framework laid out in Section 3.

<sup>21</sup> Data limitations prevent U.S. from identifying individuals in the armed forces in Australia.



- <sup>22</sup> In Australia, the samples are restricted to individuals aged 20 to 54 in 1981 and to those aged 30 to 64 in 1991 because of data limitations.
- <sup>23</sup> Agricultural industries include agricultural production and services, forestry, fishing, hunting and trapping.
- <sup>24</sup> To account for aging of the cohorts of workers, unlike in the regression data, we restrict attention to individuals aged 18–54 (20–54 in Australia) in all years.
- <sup>25</sup> For ease of exposition, we will refer to particular immigrant cohorts using the year intervals that pertain to the Australian and Canadian data, with the implied understanding that in the U.S.U.S. data the actual cohort intervals begin and end one year earlier.
- <sup>26</sup> See Aronson (1991) and Parker (2004) for reviews.
- <sup>27</sup> Examples include Chiswick (1978), Borjas (1985, 1995), Beach and Worswick (1993) and Green and Green (1995).
- <sup>28</sup> In essence, the period effect is estimated from natives, and this information is used to identify cohort and assimilation effects for immigrants.
- <sup>29</sup> The full set of regression results are available from the authors upon request.
- <sup>30</sup> In Australia, we restrict the sample to respondents who worked 16 (15) or more hours in their main job in the reference week in 1981 (1991).
- <sup>31</sup> In Australia, all variables pertain to the reference week, therefore, we continue to define self-employment as described in Section 4.
- <sup>32</sup> The results reported here include those with self-employment income comprising 25 percent or more of the previous year's earnings. However, the results are not sensitive to this cut-off—in part because the incidence of individuals with multiple sources of income is relatively rare.
- <sup>33</sup> Unfortunately, the Australian census does not distinguish an individual's earnings from his other sources of income, so for Australia we are forced to use weekly personal income as our proxy for wages.
- <sup>34</sup> While the income measure in Australia is categorical, we convert it into a continuous variable by assigning each individual the midpoint of his income category. To ensure this does not effect our results, we estimate the model using both interval and censored regression. The results are similar and available upon request.

## REFERENCES

- Aldrich, H., Cater, J., Jones, T., McEvoy, D. and Velleman, P. (1985). Ethnic Residential Concentration and the Protected Market Hypothesis. *Social Forces* 63: 996–1009.
- Antecol, H. (2004). The History of Selective Immigration in Australia, Canada, and the United States. *Brussels Economic Review, Special Edition: Selective Immigration—Consequences for Source and Destination Countries*.
- Antecol, H., Cobb-Clark, D. A. and Trejo, S. J. (2003a). Immigration Policy and the Skills of Immigrants to Australia, Canada, and the United States. *Journal of Human Resources* 38 (1): 192–218.
- Antecol, H., Cobb-Clark, D. A. and Trejo, S. J. (2003b). The Skills of Female Immigrants to Australia, Canada, and the United States. In *Host Societies and the Reception of Immigrants*, J. G. Reitz (Ed.), pp. 327–359. Center for Comparative Immigration Studies, San Diego, CA.
- Antecol, H., Kuhn, P. and Trejo, S. (2003). *Assimilation via Prices or Quantities? Labor Market Institutions and Immigrant Earnings Growth in Australia, Canada, and the United States*. Discussion Paper No. 802, IZA, Bonn.

- Aronson, R. L. (1991). *Self-Employment: A Labor Market Perspective*. ILR Press, Ithaca, NY.
- Baker, M. and Benjamin, D. (1994). The Performance of Immigrants in the Canadian Labor Market. *Journal of Labor Economics* 12: 369–405.
- Bloom, D., Grenier, G. and Gunderson, M. (1995). The Changing Labor Market Position of Canadian Immigrants. *Canadian Journal of Economics* 28: 987–1005.
- Borjas, G. (1985). Assimilation, Changes in Cohort Quality and the Earnings of Immigrants. *Journal of Labor Economics* 3: 463–489.
- Borjas, G. (1986). The Self-Employment Experience of Immigrants. *Journal of Human Resources* 21: 485–506.
- Borjas, G. (1993). Immigration Policy, National Origin and Immigrant Skills: A Comparison of Canada and the United States. In *Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States*, D. Card and R. Freeman (Eds). University of Chicago Press, Chicago.
- Borjas, G. (1995). Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s? *Journal of Labor Economics* 13: 201–245.
- Borjas, G. and Bronars, S. (1989). Consumer Discrimination and Self-Employment. *Journal of Political Economy* 97: 581–605.
- Boyd, R. L. (1990). Black and Asian Self-Employment in Large Metropolitan Areas: A Comparative Analysis. *Social Problems* 37: 258–269.
- Briggs, V. M., Jr. (1984). *Immigration Policy and the American Labor Force*. Johns Hopkins University Press, Baltimore.
- Carroll, G. R. and Mosakowski, E. (1987). The Career Dynamics of Self-Employment. *Administrative Sciences Quarterly* 32: 570–589.
- Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-Born Men. *Journal of Political Economy* 86: 897–921.
- Chiswick, B. R. (1986). Immigration Policy, Source Countries and Immigrant Skills: Australia, Canada and the United States. In *The Economics of Immigration: Proceedings of the Conference held at the Australian National University*. Australian Government Printing Service, Canberra.
- Clark, K. and Drinkwater, S. (1998). Ethnicity and Self-Employment in Britain. *Oxford Bulletin of Economics and Statistics* 60: 383–407.
- Clark, K. and Drinkwater, S. (2000). Pushed Out or Pulled In? Self-Employment among Ethnic Minorities in England and Wales. *Labour Economics* 7: 603–628.
- Clark, K. and Drinkwater, S. (2002). Enclaves, Neighborhood Effects and Employment Outcomes: Ethnic Minorities in England and Wales. *Journal of Population Economics* 15: 5–29.
- Fairlie, R. and Meyer, B. (1996). Ethnic and Racial Self-Employment Differences and Possible Explanations. *Journal of Human Resources* 31: 757–793.
- Fairlie, R. and Woodruff, C. (2004). *Mexican Entrepreneurship: A Comparison of Self-employment in Mexico and the United States*. Working Paper, University of California Santa Cruz.
- Flota, C. and Mora, M. T. (2001). The Earnings of Self-Employed Mexican-Americans along the U.S.–Mexico Border. *Annals of Regional Science* 35: 483–499.
- Green, A. G. (1995). A Comparison of Canadian and U.S. Immigration Policy in the Twentieth Century. In *Diminishing Returns*, D. DeVoretz (Ed.). CD Howe Institute and Laurier Institution, Toronto.
- Green, A. G. and Green, D. A. (1995). Canadian Immigration Policy: The Effectiveness of the Point System and Other Instruments. *Canadian Journal of Economics* 28: 1006–1041.
- Lalonde, R. J. and Topel, R. (1992). The Assimilation of Immigrants in the U.S. Labor Market. In *Immigration and the Workforce*, G. J. Borjas and R. Freeman (Eds). University of Chicago Press, Chicago.

- Le, A. T. (2000). The Determinants of Immigrant Self-Employment in Australia. *International Migration Review* 13: 183–214.
- Lester, L. and Richardson, S. (2004). *A Comparison of Australian and Canadian Immigration Policies and Labour Market Outcomes*. Report to the Department of Immigration and Multicultural and Indigenous Affairs. Commonwealth of Australia.
- Li, P. (1998). *The Chinese in Canada*. Second Edition Oxford University Press, Toronto.
- Light, I. (1972). *Ethnic Entrepreneurs in America*. University of California Press, Berkeley, CA.
- Light, I. (1984). Immigrant and Ethnic Enterprise in North America. *Ethnic and Racial Studies* 7: 195–216.
- Light, I. and Bonacich, E. (1988). *Immigrant Entrepreneurs: Koreans in Los Angeles 1965–1982*. University of California Press, Berkeley, CA.
- Light, I. and Rosenstein, C. (1995). *Race, Ethnicity and Entrepreneurship in Urban America*. Aldine de Gruyter, New York.
- Lofstrom, M. (2002). Labor Market Assimilation and the Self-Employment Decision of Immigrant Entrepreneurs. *Journal of Population Economics* 15: 83–114.
- Mata, F. and Pendakur, R. (1998). *Immigration, Labor Force Integration and the Pursuit of Self-Employment*. RIIM Working Paper #98-05.
- Metcalf, H., Mohood, T. and Virdee, S. (1996). *Asian Self-Employment: The Interaction of Culture and Economics*. Policy Studies Institute, London.
- Miller, P. W. (1999). Immigration Policy and Immigrant Quality: The Australian Points System. *American Economic Review* 89: 192–197.
- Min, P. G. (1984). From White Collar Occupations to Small Business: Korean Immigrants' Occupational Adjustment. *The Sociological Quarterly* 25: 333–352.
- Min, P. G. (1988). *Ethnic Business Enterprise: Korean Small Business in Atlanta*. Center for Migration Studies, New York.
- Moore, R. L. (1983). Employer Discrimination: Evidence from Self-Employed Workers. *Review of Economics and Statistics* 65: 496–501.
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Phizacklea, A. (1988). Entrepreneurship, Ethnicity and Gender. In *Enterprising Women*, S. Westwood and P. Bhachu (Eds). Routledge, London.
- Rafiq, M. (1992). Ethnicity and Enterprise: A Comparison of Muslim and Non-Muslim Owned Asian Businesses in Britain. *New Community* 19: 43–60.
- Razin, E. and Longlois, A. (1996). Metropolitan Characteristics and Entrepreneurship Among Immigrants and Ethnic Groups in Canada. *International Migration Review* 30: 703–727.
- Schuetze, H. J. (2000). Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-U.S. Comparison. *Labor Economics* 7: 507–544.
- Schuetze, H. J. (2005). *The Self-Employment Experience of Immigrants to Canada*. Working Paper, University of Victoria, BC, Canada.
- Schuetze, H. J. and Bruce, D. (2004). Tax Policy and Entrepreneurship. *Swedish Economic Policy Review* 11 (2): 233–265.
- Sowell, T. (1981). *Markets and Minorities*. Basic Books, New York.
- Wong, L. and Ng, M. (2002). The Emergence of Small Transnational Enterprise in Vancouver: The Case of Chinese Immigrant Entrepreneurs. *International Journal of Urban and Regional Research* 26 (3): 508–530.
- Yuengert, A. M. (1995). Testing Hypothesis of Immigrant Self-Employment. *Journal of Human Resources* 30: 194–204.

DAVID B. AUDRETSCH<sup>a</sup> and ERIK LEHMANN<sup>b</sup>

<sup>a</sup>*Indiana University and Max Planck Institute of Economics, Entrepreneurship, Growth and Public Policy, Kahlaische Straße 10, D-07745 Jena, Germany*

*E-mail: audretsch@econ.mpg.de*

<sup>b</sup>*University of Augsburg, Chair of Management and Organisation, Universitätsstraße 16, D-86159 Augsburg, Germany*

*E-mail: Erik.Lehmann@wiwi.uni-augsburg.de*

## 6. Location and New Venture Creation

### 1. INTRODUCTION

In proposing a new theory of economic geography, Krugman (1991, p. 5) asks, “What is the most striking feature of the geography of economic activity? The short answer is surely concentration [...] production [...] is remarkably concentrated in space.” As for other fields of economics, the impact of geography has not escaped the attention of scholars of entrepreneurship. A recent wave of studies has focused on the location decision of new-firm startups. Indeed, an important finding of this literature is that the impact of geographic characteristics on location choice is anything but neutral. For example, the collection of European country studies included in the special issue of *Regional Studies* on “Regional Variations in New Firm Formation” (Reynolds, Storey and Westhead, 1994) identified a number of geographic specific characteristics that impact the location of new firms. These characteristics were generally based on those factors identified in earlier studies by Carlton (1983) and Bartik (1985).

In the Section 2 of this chapter, we explain how and why location proximity should generate benefits to knowledge based new ventures. In Section 3, the different types of knowledge outputs and different mechanisms used by new ventures to access knowledge spillovers from universities are discussed. Not only are the types of knowledge and spillover mechanisms heterogeneous, but the capacity to generate knowledge spillovers also varies considerably across universities. However, none of these studies focused on the role of accessing knowledge spillovers in the location choice decision of new firms. This oversight is surprising given that the growing literature

on technology management and the economics of innovation has found that knowledge spillovers play an important role in fostering entrepreneurship and innovative activity (Sorenson and Audia, 2000; Baum and Sorenson, 2003). In addition, spillovers from universities, research parks, as well as from private firms, have been identified as a key source promoting firm innovation and performance (Stuart and Sorenson, 2003; Link and Scott, 2005).

Thus, in the Section 4 chapter we identify the role that location plays in influencing new venture creation as well as new venture performance. In particular, we introduce a new data base consisting of 281 publicly listed new ventures in German high technology and knowledge industries is used to identify empirically in Section 5 how location choice varies for different types of knowledge and spillover mechanisms. In the last section, a summary and conclusion are provided. In particular, the evidence suggests that, in general, knowledge and technological-based new ventures have a high propensity to locate close to universities, presumably in order to access knowledge spillovers. However, the exact role that geographic proximity plays is shaped by the two factors examined in this paper—the particular knowledge context, and the specific type of spillover mechanism.

## 2. NEW VENTURES

Within the economics literature, the prevalent theoretical framework has been the general model of income choice. The model of income choice dates back at least to Knight (1921), but was more recently extended and updated by Lucas (1978), Kihlstrom and Laffont (1979), Holmes and Schmitz (1990) and Jovanovic (1994), and addresses the fundamental question, “*Why and how do individual economic agents decide to start a new venture?*” Thus, the unit of analysis is at the level of the individual economic agent. In its most basic rendition, individuals are confronted with a choice of earning their income either from wages earned through employment in an incumbent enterprise or else from profits accrued by starting a new venture. The essence of the income choice is made by comparing the wage an individual expects to earn through employment,  $W^*$ , with the profits that are expected to accrue from a new-venture,  $P^*$  (Parker, 2003, 2004). Thus, the probability of starting a new venture,  $Pr(s)$ , can be represented as:

$$Pr(s) = f(P^* - W^*). \quad (1)$$

The model of income choice has been extended by Kihlstrom and Laffont (1979) to incorporate aversion to risk, by Lazear (2002) to include characteristics of human capital and by Lucas (1978) and Jovanovic (1994) to

explain why firms of varying size exist and has served as the basis for empirical studies of the decision to start a new venture in a broad range of countries, time periods and contexts (Audretsch, 2003).

This view of entrepreneurship corresponds to that in a different scholarly tradition—management—provided by Gartner and Carter (2003): “Entrepreneurial behavior involves the activities of individuals who are associated with creating new organizations rather than the activities of individuals who are involved with maintaining or changing the operations of on-going established organizations.”

Both the field of management and psychology have provided insights into the decision process leading individuals to start a new venture. This research trajectory focuses on the emergence and evolution of entrepreneurial cognition. Stevenson and Jarillo (1990) assume that entrepreneurship is an orientation toward opportunity recognition. Central to this research agenda are the questions, “How do entrepreneurs perceive opportunities and how do these opportunities manifest themselves as being credible versus being an illusion?” Kruger (2003) examines the nature of entrepreneurial thinking and the cognitive process associated with opportunity identification and the decision to undertake entrepreneurial action. The focal point of this research is on the cognitive process identifying the entrepreneurial opportunity along with the decision to start a new venture. Thus, a perceived opportunity and intent to pursue that opportunity are the necessary and sufficient conditions for entrepreneurial activity to take place. The perception of an opportunity is shaped by a sense of the anticipated rewards accruing from and costs of becoming an entrepreneur. Some of the research focuses on the role of personal attitudes and characteristics such as self-efficacy (the individual’s sense of competence), collective efficacy and social norms. Shane (2000) has identified how prior experience and the ability to apply specific skills influence the perception of future opportunities.

The concept of the entrepreneurial decision resulting from the cognitive processes of opportunity recognition and ensuing action is introduced by Shane and Eckhardt (2003) and Shane and Venkataraman (2001). They suggest that an equilibrium view of entrepreneurship stems from the assumption of perfect information. By contrast, imperfect information generates divergences in perceived opportunities across different people. The sources of heterogeneity across individuals include different access to information as well cognitive abilities, psychological differences and access to financial and social capital.

One of the best data sources available to analyze the cognitive process triggering the entrepreneurial decision is provided by the Panel Study of Entrepreneurial Dynamics (PSED) which consists of a longitudinal survey study on 830 individuals that were identified while they were in the process of starting a new business. The unique feature of the database is that it provides information on how the entrepreneurial opportunity and action was

conceived and operationalized (Gartner and Carter, 2003). Kim, Aldrich and Keister (2003) use the PSED to test the theory that access to resources, in the form of financial resources such as household income and wealth, and human capital, in the form of education, prior work experience, entrepreneurial experience and influence from family and friends, affect the decision to become an entrepreneur.

As the Kim, Aldrich and Keister (2003) paper suggests, the external environment has been found to strongly influence the entrepreneurial decision. The greatest focus of research has been on the influence of networks on the cognitive process involving entrepreneurship. Thornton and Flynn (2003) argue that geographic proximity leads to networking, which both creates opportunities as well as the capacity to recognize and act on those opportunities. They suggest that networks in which trust is fostered involve a context facilitating the transmission of tacit knowledge. In comparing Route 128 around Boston with Silicon Valley, Saxenian (1994) documented how entrepreneurial advantages are based on differences in network structures and social capital.

Research has considered both the formation as well as the impact of networks on entrepreneurship. Hoang and Antoncic (2001) characterize research as systematically focusing on network content, network governance and network structure. Thus, there is considerable evidence and theory suggesting that external linkages and influences will shape the context of the entrepreneurial decision made by the individual.

Accordingly, there is a solid research tradition focusing on the decision confronting individuals to start a venture. Theory and empirical evidence provide compelling reasons to conclude that both characteristics specific to the individual as well as context external to the individual help shape the cognitive processes guiding the entrepreneurial decision.

Recognition of the role that firm-specific knowledge investments could play in accessing and absorbing external knowledge, and therefore enhancing the innovative output of the firm, triggered an explosion of studies focusing on potential sources of knowledge that are external to the firm. Some studies examined the role of licensing, cooperative agreements and strategic partnerships, all of which involve a formal agreement and a market transaction for the sale of knowledge. Thus, these all represent mechanisms by which a firm can access knowledge produced by another firm. As Cohen and Levinthal (1989) emphasized, presumably internal investment in knowledge is a prerequisite for absorbing such external knowledge—even if can be accessed.

### 3. THE ROLE OF LOCATION

Geography and spatial location also influence entrepreneurship. The important roles that geographic clusters and networks play as a determinant

of entrepreneurial activity was identified in Europe and only recently has been discovered within the North American context (Porter, 1990, 2000; Saxenien, 1994). By contrast, there is a longer and richer tradition of research linking entrepreneurship to spatial clusters and networks in Europe. However, most of these studies have been in social science fields other than economics. For example, Becattini (1990) and Brusco (1990) identified the key role that spatial clusters and networks play in promoting SMEs in Italy. While such networks and clusters were generally overlooked or ignored in North America, with publication of Saxenien's book, *Regional Advantage* (1994), which documented how spatial networks generated entrepreneurial activity in Silicon Valley and Route 128 around Boston, it became clear and accepted that spatial agglomerations were also important in the North American context.

An important distinction between the European literature and studies and the emerging literature in North America was the emphasis on high technology and knowledge spillovers in the North American context. By contrast, the European tradition focused much more on the role of networks and clusters in fostering the viability of SMEs in traditional industries such as textiles, apparel and metalworking. For example, seminal studies by Becattini (1990) and Brusco (1990) argue that small and new ventures enjoy a high degree of stability when supported by networks in Italy. A rich literature has provided a compelling body of case studies spanning the textile industries of northern Italy to the metal working firms of Baden Wuerttemberg (Piore and Sabel, 1984), documenting the long-term viability and stability of small and new firms embedded in the so-called industrial districts of Europe. Pyke and Sengenberger (1990) argue that through the support of an industrial district, small firms in European spatial clusters have been able to compensate for what would otherwise be an inherent size disadvantage. According to Pyke and Sengenberger (1990, an industrial district is a geographically defined production system, involving a large number of enterprises engaging in production at a wide range of stages and typically involved in the production of a homogeneous product. A particularly significant feature of Italian industrial districts is that almost all of the firms are small or even micro-enterprises. Examples of such industrial districts include Prato, Biella, Carpi and Castelfreddo, which specialize in textile (coolants in Castelfreddo); Vigevano, Montebellune and Montegranaro where shoes are manufactured (ski boots in Montebellune); Pesaro and Nogara which manufacture wooden furniture; Sassuolo where ceramic tiles are produced.

Brusco (1990) emphasizes the cooperation among network firms within an industrial district. Such cooperation presumably reduces any size-inherent disadvantages and improves the viability of small firms operating within the network. According to Pyke and Sengenberger (1990, p. 2), "A characteristic of the industrial district is that it should be conceived as a social and economic whole. That is to say, there are close inter-relationships between the different



social, political and economic spheres, and the functioning of one, say the economic, is shaped by functioning and organization of the others.” Grabher (1993) similarly argues that the social structure underlying industrial networks contributes to the viability of small firms that would otherwise be vulnerable if they were operating in an isolated context.

A different research trajectory focused on flows of knowledge across firms where no market transaction or formal agreement occurred, or what has become known as knowledge spillovers. The distinction between knowledge spillovers and technology transfer is that in the latter a market transaction occurs, whereas in the case of spillovers the benefits are accrued without an economic transaction.

While Krugman (1991) and others certainly did not dispute the existence or importance of knowledge spillovers, they contested the claim that knowledge spillovers should be geographically bounded. Their point was that when the marginal cost of transmitting information across geographic space approaches zero, there is no reason to think that the transmission of knowledge across geographic space should stop simply because it reaches the political border of a country, city, state or country.

However, von Hippel (1994) explained how knowledge is distinct from information and requires geographic proximity in transmitting ideas that are highly dependent upon their context, inherently tacit and have a high degree of uncertainty. This followed from Arrow (1962), who distinguished economic knowledge from other economic factors as being inherently nonrival in nature, so that knowledge developed for any particular application can easily spill over to generate economic value in very different applications. As Glaeser, Kallal, Scheinkman and Shleifer (1992, p. 1126) have observed, “Intellectual breakthroughs must cross hallways and streets more easily than oceans and continents.”

Thus, a distinct research trajectory developed in the late 1980s and early 1990s trying to identify the impact of location on the innovative output of firms. These studies addressed the question “Holding firm-specific knowledge inputs constant, is the innovative output greater if the firm is located in a region with high investments in knowledge?” The answer to this question was provided in a series of studies shifting the unit of observation for testing the model of the knowledge production function from the firm to a spatial unit of observation, such as a city, region or state.

Studies identifying both the extent but also the localization of knowledge spillovers were also based on the model of the knowledge production function. Jaffe (1989) modified the knowledge production function approach to a model specified for spatial and product dimensions:

$$I_{si} = IRD^{\beta_1} \times UR_{si}^{\beta_2} \times (UR_{si} \times GC_{si}^{\beta_3}) \times \varepsilon_{si}, \quad (2)$$

where  $I$  is innovative output,  $IRD$  is private corporate expenditures on R&D,  $UR$  is the research expenditures undertaken at universities, and  $GC$  measures the geographic coincidence of university and corporate research. The unit of observation for estimation was at the spatial level,  $s$ , a state and industry level  $i$ . Estimation of equation (1) essentially shifted the model of the knowledge production function from the unit of observation of a firm to that of a geographic unit.

Compelling and consistent evidence provided first by Jaffe (1989) but later confirmed by Acs, Audretsch and Feldman (1991 and 1994), Feldman (1994), Jaffe, Trajtenberg and Henderson (1993) and Audretsch and Feldman (1996) suggested that, in fact, the presence of external knowledge sources in geographically bounded regions increased the innovative output of firms located in those regions. Thus, there was clear and compelling econometric evidence suggesting that external investments in geographically bounded regions would yield an increased level of innovative output by the firms located in that region as a result of knowledge spillovers.

The new findings from the studies on spatially bounded knowledge spillovers, in two main ways, supported the knowledge production model of firm innovation. First, the firms were still assumed to be exogenous, and second, knowledge inputs were still found to be important determinants of innovative output. The main distinction lies in the unit of analysis. Because of knowledge spillovers, the link between knowledge inputs and firm innovative output was found to be more important for spatial units of observation than at the level of the firm.

## 4. UNIVERSITY SPILLOVERS

### *4.1. Geographical Proximity, Spillovers and New Venture Formation*

As the previous section concludes, a basic tenet in the literature is that university spillovers lower the costs of firms to accessing and absorbing knowledge spillovers. This leads to the formulation of the first hypothesis.

*Hypothesis H1: The location strategy of new ventures involving geographic proximity to an university is more important when the research level of the university is high.*

Such spillovers could be transmitted through certain conduits across geographic space such as the channels of communication, the social system, or a kind of technology diffusion process. Most of those benefits could not be obtained by markets or ensured by contractual arrangements because much of the tacit knowledge is transferred via communication channels. According to the theory of communication, those channels can be decomposed into two

categories: Communication transmitted via articles as a means of mass media in the scientific world and interpersonal communication. Such interpersonal communications are important influences in determining the speed and thus the costs of the diffusion of knowledge.

The primary conduit to achieve and absorb the spillover effects is through strategic location yielding geographical proximity to the knowledge source. Thus, the new venture location decision should be influenced by the activities of the local universities. In particular, the role that geographic proximity to the university plays in accessing spillovers should be shaped by the relative importance of the transmission of codified knowledge through the mass media (in the scientific context) versus the relative importance of tacit knowledge.

The relative importance of codified knowledge is reflected by the predominance of articles published in high-quality scientific journals—the mass media channel. By contrast, the relative importance of tacit knowledge is reflected by the number of fresh graduates from the university, which serves as a measure for the intense demand for labor and interpersonal communication.

The second hypothesis is therefore based on the anticipated impact of the relative importance of tacit knowledge versus codified knowledge on the location benefits of geographic proximity to an university.

*Hypothesis H2: The strategic advantage bestowed by new venture geographic proximity to an university will be greater where the university generates research output with a high tacit knowledge content. By contrast, universities generating research output with a low tacit knowledge content offer less of a strategic advantage to new venture location proximity.*

Strict adherence to the scientific method assures that academic publications embody a high component of codified and specific knowledge in the natural sciences (Stephan, 1996). By contrast, with a more limited applicability of the scientific method, publications in the social sciences embodied less codified knowledge (Stephan, 1996). However, only a small field in the social science, like economics and econometric theory, contain specific and codified articles (see Audretsch, Lehmann and Warning, 2005a, b). The distinction between codified and less codified research articles also reflects the degree of specific knowledge. Research fields in the natural science are to a greater extent specific to certain industries as fields in social science. As an example, the role of social and human capital is not only specific to one industry as it is in the case of biochemistry or medicine.

In contrast to other studies, we are able to decompose the output of academic research into natural science and social science. Prominent examples of research spillover effects in science are demonstrated in the case of biotechnology by Audretsch and Stephan (1996, 1999), Zucker, Darby and

Armstrong (1998) or, more general in the case of patents or investments in R&D (Jaffe, 1989). To our understanding, there is no empirical study which primarily focuses on research spillover effects in social science. We assume that such effects are less special and firm specific than those in science. For example, seminars, presentations and conferences in accounting, finance or management are valuable for every firm, independently from the type of production or the industry classification. Those effects could be measured by the number of articles published in this field. Since spillover effects in social science are relevant for all firms we expect no differences across industries, compared to research spillover effects in science.

*Hypothesis H3: In the natural sciences codification inherent in published articles enables absorption over a longer geographic distance. In the social sciences geographic proximity is required to absorb knowledge spillovers.*

Spatial proximity to universities can also generate positive externalities that can be accessed by the new venture through hiring fresh graduates. First, fresh graduates may be important channels for disseminating knowledge from academia to the local high technology industry (Varga, 2000). Other externalities may rise through the close location *per se*. Local proximity lowers the search costs for both new ventures and graduates. This may lead to some competitive advantage over similar new ventures which are not located close to universities, especially when high skilled labor is a scarce resource and there is intense competition about high potentials.<sup>1</sup>

*Hypothesis H4: The greater the output of student graduates, the lower is the distance between the new venture and the closest university.*

A similar logic as for research activities holds for the relative components of tacit and specific knowledge embodied in graduates. Graduates in the natural sciences presumably embody a higher component of human capital specific to a particular science and technology. The knowledge of biologists, information engineers, physics or chemists is more specific to a particular firm and industry compared to the knowledge of economists, sociologists and graduates in business. Graduates embody specific knowledge in the natural sciences, leading them to locate within geographic proximity to the university.

However, since we focus on young and innovative new ventures, there are some restrictions compared to past research. Since those new ventures are small and constraint in their financial resources, they are less able to act as research partners in natural or physical science for universities by providing funds or physical assets. Also academic research in those areas are not always a kind of public good if academic researchers compete with researchers from new ventures which are not included in the research relationship. Thus, research

spillover cannot easily be exploited. The main channel to participate from the academic research is to employ graduates from local universities. By contrast, graduates embody general knowledge in the social sciences, leading them to more diffused locations. Thus, the fifth hypothesis predicts that the ability of new ventures to access and absorb social science graduates is less dependent upon specific location than is the case for natural science graduates.

*Hypothesis H5: A new venture's ability to access and absorb graduates is less dependent upon geographic space in the social sciences than in the natural sciences.*

The preceding five hypotheses focus on the strategic location decision confronting the new venture and how it will vary according to different knowledge conditions, types of university research output, and the relative importance of tacit and codified knowledge. New venture performance should also be influenced by location. In particular, access to knowledge spillovers should generate a superior performance. Measuring performance in early stage knowledge-based technology new ventures is well known to be difficult since traditional performance measures, such as profits, do not apply (see Audretsch and Lehmann, 2005). Thus, we take the duration from new venture foundation until the listing on the stock market as a performance measure (see also Stuart, Hoang and Hybels, 1999; Stuart and Sorenson, 2003). The availability of external equity is an important resource for high-tech new ventures. Because new ventures lack existing cash flow to finance investment and future growth, the time window from foundation to the listing on the stock market is important for the efforts of the ventures to exploit new technologies. Since high-tech new ventures are associated with a high risk of default and asymmetric information, they suffer from credit restrictions by banks (Audretsch and Lehmann, 2004). The only way to receive financial resources to grow is equity, provided by venture capitalists or large firms. The vast amount of capital, however, is reached when the new ventures are able to go public and disseminate their shares on the stock market. This leads to the final hypothesis.

*Hypothesis H6: The greater the geographic proximity of the new venture to the university, the lower should be the time window from new venture foundation to the stock market listing.*

## 4.2. Methodology

*4.2.1. Sample* To test the hypotheses that venture foundation depends on geographical proximity and university spillovers we use a unique dataset of all of the German new ventures listed on the *Neuer Markt*. The total population of new ventures listed on the *Neuer Markt*, Germany's counterpart of the *NASDAQ*, between 1997 and 2002 was 295. This dataset consists of all 295

publicly listed German new ventures and was collected combining individual data from IPO prospectuses along with publicly available information from on-line data sources including the *Deutsche Boerse AG* [www.deutsche-boerse.com](http://www.deutsche-boerse.com). We use this database for several reasons. First, the included ventures include highly innovative industries like biotechnology, medical devices, life sciences, e-commerce and other high-technology industries which represent the knowledge-based economy. Second, studies from the U.S. provide strong evidence for the growth effect of clusters influenced by the presence of research active university (Feldman, 2000). This dataset enables us to follow this line of research. Third, this data set represents the technological change in the German Business sector from the predominance of medium-sized firms in the production and manufacturing toward the high-technology and service sector. Finally, in Germany, such data are not available for privately held new ventures.<sup>2</sup> We complete this dataset by adding university-specific variables which are individually collected from the universities and the research database from the ISI (Information Sciences Institutes). We did not include research institutes since they only have a few graduates.

*4.2.2. Variables and Measurement* To test the six hypotheses posed in the previous section, we use two different dependent variables. First, we take the DISTANCE to the closest university. Since universities in Germany are more geographically concentrated compared to the U.S., we need a measure which is sensible enough for small variations. The distance is measured in kilometers using the online database of the *German Automobile Club* ([www.adac.de](http://www.adac.de)). All ventures located within a radius of 1.5 kilometers are classified as belonging in the distance category of 1 kilometer.

The second endogenous variable is CLUSTER. This ordinal variable captures geographic proximity by focusing on the location closest to the university, within the same city and outside this area. This measurement is analogous to SMSAs and similar measures indicating location proximity. The variable CLUSTER takes the value of one if the new ventures is located within a close radius of 8 kilometers (the median value) around the university. If the firm is located within a radius of 20 kilometers the variable takes on the value of 2, or 3 if the location is outside the radius of 20 kilometers.

We include the DURATION or time-to-IPO as the third endogenous variable. Since we have no exact information about the month of firm formation, we used the number of years to measure the time dimension.

The independent and predictor variables are as follows. To measure research output of universities we include the number of articles published in high quality journals (see Zucker, Darby and Armstrong, 1998; Audretsch and Stephan, 1996) and the number of graduates in 1997. Since university spillovers are not restricted to patented inventions and occur solely in the natural sciences,

we include measures for social science research output as well as natural science output. This enables us to discriminate between the sources of spillover effects. Since knowledge-based industries include services such as media and entertainment, service or e-commerce, spillovers can also be generated by fields without high patent activities. Articles published in social science (*SSCI*) are measured by the ISI-database *SSCI* (Social Science Citation Index). Articles in natural sciences (*SCI*) are taken from the *SCI* (Science Citation Index). We included the number of listed papers for each university published from 1993 until 2000 (see Warning, 2004).<sup>3</sup> Furthermore, the location decision may also be influenced by clustering effects of universities. Thus, we include the number of universities located in a city to capture this effect (*CLUSTER*).

Although a number of studies provide strong evidence that the number and quality of articles published in high quality journals influences the location of innovative activity, Varga (2000) points to the effects of new graduates as an important mechanism for transmitting the latest knowledge from academia to firms located in the same geographic area. However, spillover effects may not only arise by the knowledge transmission of students but also by their employment effects. The nearby location enables firms to attract high skilled employees with lower costs. Recent graduates also have the possibility to work at local companies without leaving their social network. We further control for graduates from science (*SCIGRADS*) and from social science (*SSCIGRADS*). Both measures are from the year 1997.

As pointed out by McWilliams and Siegel (2000), among others, a major determinant of the performance of highly innovative firms is their spendings in R&D. To capture this effect, we include a firm's investment in R&D. However, the observation of a value of zero investments in R&D could be either because the investment is negligible or the case of nonreporting. Thus, we include a dummy variable (*R&DREPORT*) when R&D are missing to control for the possibility that nonreporting firms are discretely different from reporting firms. Simply eliminating observations with missing values for these variables is undesirable because it significantly reduces the sample size and biases the sample in favor of R&D-intensive firms.

Finally, previous research has shown that spillover effects differ between industries in their necessity and capability to absorb spillover effects (Jaffe, 1989; Cohen and Levinthal, 1990; Henderson and Cockburn, 1994). To control for specific industry effects, we include dummy variables for the following industries: Software, E-Services, E-Commerce, Computer & Hardware, Telecommunication, Biotechnology, Medicine & Life Science, Media & Entertainment and High-Technology. In addition, to control for the impact of the life cycle of the firm (Agarwal, Echambadi and Sarkar, 2002), we include firm age (*AGE*).

### 4.3. Analysis and Methodology

To test the six hypotheses raised in Section 3.1, we employ three different empirical estimation methods: negative binomial regressions, ordered probit and hazard models.

First, we employ the negative binomial regression model as the analytical technique for estimating the impact of university research output on the strategic location choice of firms. The underlying assumption is that distance as measured in kilometers could be interpreted as count data. Since ordinary least squares regression is inappropriate for the count dependent variables that have large numbers of the smallest observation and remaining observations taking the form of small positive numbers, Poisson-regressions seems to be more appropriate (Greene, 2003). However, the assumption for a Poisson regression, the equality of mean and variance of the exogenous variable, is rejected by several tests. Thus, we apply the negative binomial regression model to overcome this problem of “over-dispersion.” Also, this statistical technique is designed for maximum likelihood estimation of the number of occurrence of nonnegative counts like the event of location.

Second, we apply ordered probit estimation as a robust test for the negative binomial regression. According to studies which take SMSAs or related areas as the measure for geographic proximity, we use the ordinal variable CLUSTER as the dependent variable. The regression is then based on the maximum-likelihood method. In place of the traditional calculation we estimate the regression with the Huber/White/sandwich estimator of variance.

Finally, to make an inference about the impact of location strategy on firm performance, we estimate a simple Cox proportional hazard approach to measure the duration effect from firm foundation to the listing on the stock market. The probability for being listed within an interval  $(0, t)$  is given by the distribution function  $F(t)$ . The derivation is called density of  $T$  and named by  $f(t)$ . The complement of the distribution function is called survivor function  $S(t) = 1 - F(t)$  and indicates the probability for not being listed on the stock market at time  $t$ . A central element in the analysis of duration data is the concept of the hazard function. It is defined as the conditional probability for being listed on the market within the interval  $t + \Delta t$  given the firm has not been listed on the market at time  $t$  (see Kiefer, 1988).

## 5. EMPIRICAL RESULTS

### 5.1. Descriptive Statistics

Table 6-1 presents some descriptive statistics of both dependent and independent variables. The closest location is one kilometer and the maximum



TABLE 6-1 *Descriptive statistics*

|            | Mean      | Std. Dev. | Min | Max    | 25%Cent | Median | 75%Cent |
|------------|-----------|-----------|-----|--------|---------|--------|---------|
| Distance   | 16.69     | 23.457    | 1   | 177    | 1       | 7      | 21      |
| SCI        | 10,689.17 | 11,947.55 | 8   | 34,148 | 169     | 6,357  | 13,742  |
| SSCI       | 596.41    | 607.06    | 0   | 1,694  | 98      | 491    | 816     |
| SCI-Grads  | 20,494    | 15,292.45 | 0   | 47,112 | 4936    | 7,725  | 9,395   |
| SSCI-Grads | 7,270     | 3,921.09  | 0   | 20,570 | 6993    | 15,831 | 30,290  |
| Age        | 10.27     | 11.11     | 0.1 | 107    | 3       | 8      | 14.25   |

distance is 177 kilometers away from the nearest university. The skewed distribution of the data is reflected by the difference between the mean and median values. While the arithmetic mean distance is about 17 kilometers, the median shows that 50% of the firms are located within an area with the radius of 7 kilometers. The 25% (75%) centile demonstrates that 25% (75%) of the firms are located within a small radius of 1 (21) kilometer. Thus, location proximity to an university for the 295 firms in the data set is a first hint that university spillover effects may influence the strategic location decision. The descriptive statistics also show that the data is over-dispersed and thus that the alternative regression, the Poisson regression, is inappropriate.<sup>4</sup>

Table 6-1 indicates that research activities and the number of graduated students vary considerably across the universities. A comparison between the mean and median exhibits the skewed number of papers in both the social sciences and natural sciences. On average, each university published about 600 papers in social science and more than 10,000 articles in natural science. However, the number of articles published by 50% of the universities is much lower. Also the number of graduates differs across universities.<sup>5</sup>

Interestingly, the number of articles and graduate students varies not only across universities but also across the two fields. While the mean university publishes twice as many articles in the natural sciences compared to the social sciences, this difference increases with the number of published papers. While 50% of the universities publish about 500 articles in social science, there are more than 6,300 papers in science. The opposite trends can be found for the number of graduates.

The data presented in Table 6-1 show that most of the firms are strikingly young. Half of the firms in our sample are eight years old or less. Also, 25% of the firms are younger than three years.

Table 6-2 provides the correlation between the included variables. The high correlation between the articles published in SCI and SSCI demonstrates that universities are either research active—or not—independent from the discipline. Interestingly, there is a high correlation between the articles published in social science and the number of graduates in these fields. This may be due to size effects. Such effects, however, could not explain the rather low correlation

TABLE 6-2 Correlation matrix

|            | KM      | SCI    | SSCI   | SCI-Grads | SSCI-Grads |
|------------|---------|--------|--------|-----------|------------|
| SCI        | 0.0119  | 1      |        |           |            |
| SSCI       | -0.0200 | 0.9585 | 1      |           |            |
| SCI-Grads  | -0.0408 | 0.1621 | 0.0446 | 1         |            |
| SSCI-Grads | -0.0486 | 0.7620 | 0.7986 | 0.0717    | 1          |
| Age        | 0.1057  | 0.0066 | 0.0101 | 0.0784    | -0.0435    |

between the number of graduates in the natural sciences and the number of articles published in this field.

### 5.2. Firm Proximity and Research Output

Table 6-3 utilizes two different methods to estimate the distance from the location of the firm and the closest university as a function of the set of independent variables discussed above, which include the age of the firm, the social science and natural science research activities by the universities and their output of graduate students. To correct for a misspecification of the independent variable, we estimate both negative binomial and ordered probit regression models. The results for both estimations are listed in Table 6-3.

If geographic proximity is important to access and absorb knowledge spillovers, we expect a negative sign on the estimated regression coefficients, which would indicate that research and education outputs induce the founder to pursue a strategy of locate within close geographic proximity of an university. A positive sign on the estimated regression coefficient would indicate that accessing the university output is not important to the firm, or else geographic proximity is not essential to access knowledge spillovers.

As the empirical results in Table 6-3 suggest, both estimation methods provide evidence that the distance between firm location and the closest university is positively related to the university outputs of research and human capital. Both the number of articles published in social sciences, as well as the number of graduates influences the strategic decision to locate with a geographic proximity to an university. Also the number of universities is highly significant and indicates a lower and closer distance between firms and the nearest university. These results confirm hypothesis *H1*. In contrast, we find a significant and positive sign of the coefficient of the research output in the natural sciences which confirms the finding in Audretsch and Stephan (1996) that geographic proximity is not important where codified knowledge plays an important role. Thus, our results confirm hypothesis *H1* only for social science but not science.

Estimation of the models in Table 6-3 assumes that the relationships between each of the independent and the dependent variables are homogeneous

TABLE 6-3 *Negative binomial and ordered probit regressions estimating geographic proximity to an university*

|            | Negative binomial regression | Ordered probit    |
|------------|------------------------------|-------------------|
| Age        | 0.0094 (1.23)                | 0.0155 (1.98)**   |
| Cluster    | -0.3004 (3.01)***            | -0.2209 (2.26)**  |
| SSCI       | -937.9 (1.96)**              | -1723 (2.94)***   |
| SCI        | 66.7 (2.87)***               | 84.1 (3.06)***    |
| SCI-Grads  | -36.4 (1.68)*                | -44.4 (2.05)**    |
| SSCI-Grads | -3.57 (0.46)                 | 12.3 (1.50)       |
| Software   | -0.0336 (0.41)               | 0.0234 (0.26)     |
| Service    | -0.2156 (0.94)               | -0.3480 (1.45)    |
| E-commerce | 0.2086 (0.62)                | 0.1183 (0.34)     |
| Hardware   | 0.2827 (0.90)                | -0.0903 (0.29)    |
| Telecom    | -0.1703 (0.53)               | 0.1532 (0.48)     |
| Biotech    | 0.1324 (0.36)                | 0.2240 (0.63)     |
| Medtec     | -0.5708 (1.33)               | -0.6191 (1.40)    |
| Media      | -0.9633 (3.43)***            | -0.8479 (2.74)*** |
| Constant   | 3.4570 (12.21)***            | —                 |
| Pseudo R2  | 0.0246                       | 0.0913            |

The coefficients of SSCI, SCI and SSCI-Grads and SCI-Grads are multiplied with  $10^{-6}$ . Z-values are in brackets. The baseline are firms in the technology sector. The asterisks, \*, \*\*, and \*\*\* indicate significance at the 10-percent, 5-percent, and 1-percent level, respectively. The number of observations is 285.

across industries. To identify whether this assumption is true, in Table 6-4 the negative binomial model is estimated separately for each industry.

The negative coefficients in some industries of the measure of social science research output suggest that greater research output increases the importance of geographic proximity to the university.

By contrast, in the natural sciences, the positive coefficient in most industries indicates exactly the opposite—geographic proximity to an university becomes less important as research output increases. This would suggest that accessing and absorbing knowledge spillovers requires close geographic proximity in the social sciences but not in the sciences. As Stephan (1996) suggests, this may reflect a greater propensity for the scientific method to result in codified knowledge in the natural sciences than in the social sciences, where no common methodological approach has been adopted.

The negative coefficients of the number of graduates in the natural sciences (SCIGRADS) in eight of the nine knowledge-based industries suggests that, at least in most of the high-tech industries, firms choose to locate close to universities with a high yield of natural science graduates. By contrast, this holds for only one industry in the social sciences. The different impact of human capital output by the universities on the location choice of firms

TABLE 6-4 Negative binomial regressions estimating geographic proximity to an university

| Branch (N)                   | Age               | Cluster              | SSCI                  | SCI                | SCI-Grads           | SSCI-Grads         | Constant            | P-R2   |
|------------------------------|-------------------|----------------------|-----------------------|--------------------|---------------------|--------------------|---------------------|--------|
| Software (55)                | 0.0422<br>(1.95)* | -0.6782<br>(3.31)*** | 127<br>(0.10)         | 16.8<br>(0.25)     | -112.1<br>(2.31)**  | 21.4<br>(1.17)     | 3.5889<br>(6.93)*** | 0.0345 |
| E-services (69)              | 0.0195<br>(1.15)  | -0.7317<br>(3.48)*** | -404.6<br>(0.45)      | 48.1<br>(0.99)     | 7.97<br>(0.14)      | -13.1<br>(0.95)    | 3.5054<br>(5.75)*** | 0.0395 |
| E-commerce (17)              | 0.0912<br>(1.47)  | -0.5135<br>(1.19)    | -5586.1<br>(3.09)***  | 321.6<br>(3.46)*** | -338.5<br>(1.06)    | 31.3<br>(0.58)     | 4.3945<br>(4.76)*** | 0.0908 |
| Computer & Hardware (22)     | 0.0044<br>(0.21)  | 0.6940<br>(1.80)*    | -8530.0<br>(3.13)***  | 382.3<br>(2.89)*** | -234.6<br>(3.48)*** | 48.5<br>(1.11)     | 3.0857<br>(3.94)*** | 0.1185 |
| Telecommunication (20)       | 0.0009<br>(0.04)  | -0.4155<br>(1.84)*   | -2900.4<br>(1.52)     | 155.6<br>(1.79)*   | -18.8<br>(0.41)     | 15.2<br>(0.78)     | 3.7298<br>(4.99)*** | 0.0756 |
| Biotechnology (15)           | -0.0081<br>(0.14) | 0.6055<br>(1.82)*    | -12646.6<br>(3.39)*** | 445.6<br>(3.06)*** | -66.6<br>(0.75)     | 150.4<br>(2.78)*** | 2.0721<br>(2.89)*** | 0.0840 |
| Medicine & Life Science (10) | -0.0895<br>(0.86) | 0.6099<br>(1.21)     | -2044.9<br>(0.90)     | -22.5<br>(0.24)    | -118.7<br>(0.29)    | 126.2<br>(1.45)    | 0.8345<br>(0.35)    | 0.2282 |
| Media & Entertainment (38)   | -0.0331<br>(1.15) | 0.6046<br>(1.94)*    | -1298.6<br>(1.25)     | -2.74<br>(0.07)    | -51.0<br>(0.86)     | 41.5<br>(1.71)*    | 0.8395<br>(0.89)    | 0.0336 |
| High-technology (28)         | -0.0020<br>(0.19) | -0.2962<br>(0.63)    | 72.0<br>(0.04)        | 18.0<br>(0.15)     | -59.8<br>(0.47)     | 0.0314<br>(0.00)   | 3.8623<br>(2.84)*** | 0.0039 |

The coefficients SSCI, SCI, SSCI-Grads and SCI-Grads are multiplied with  $10^{-6}$ . Z-values are in brackets. P-R2 is the pseudo R2. The asterisks, \*, \*\*, and \*\*\* indicate significance at the 10-percent, 5-percent, and 1-percent level, respectively. Firms which could not be located into one of those industries are not estimated separately.

presumably reflects the high component of specific skills embodied in the natural sciences but more general skills in the social sciences. For example, students in economics, business or sociology do not really differ systematically across universities in Germany, although the research intensity in each field is different. In contrast, students in natural science differ extremely in their specialization. As an example, in Biology, graduates in botanic are no close substitutes for graduates in biotechnology. This may not hold for students in business with a specialization in either finance or accounting. Thus, the heterogeneity of the students may lead founders to locate close to the university with the expected students while students in social science are more easy available on the labor market.

### 5.3. Geographic Proximity and Firm Performance

To test the final hypothesis that the location decision impacts firm performance, we estimate a hazard model. The results are shown in Table 6-5. The sign of the estimated coefficient indicates the direction of the effect of the

TABLE 6-5 *Results from the semi-parametric Cox regression*

|            | Cox-regression   |
|------------|------------------|
| Distance   | -0.0048 (1.74)*  |
| Cluster    | 0.0747 (1.03)    |
| SSCI       | -739.1 (1.80)*   |
| SCI        | 29.3 (1.42)      |
| SCI-Grads  | 30.8 (2.05)**    |
| SSCI-Grads | 0.0709 (0.01)    |
| R&D        | 0.1138 (1.70)*   |
| R&DReport  | -0.0119 (0.90)   |
| Software   | -0.0590 (1.45)   |
| Service    | -0.1209 (0.71)   |
| E-commerce | 1.2153 (4.64)*** |
| Hardware   | -0.2767 (1.20)   |
| Telekom    | 0.1462 (0.48)    |
| Biotech    | 0.2597 (0.90)    |
| Medtec     | 0.0796 (0.32)    |
| Media      | 0.5042 (2.46)**  |
| Technology | -0.3100 (1.31)   |
| Pseudo R2  | 0.0163           |
| LL         | -1298.4183**     |

The coefficients SSCI, SCI, SSCI-Grads and SCI-Grads are multiplied with  $10^{-6}$ . Z-values are in brackets. P-R2 is the pseudo R2. The asterisks, \*, \*\*, and \*\*\* indicate significance at the 10-percent, 5-percent, and 1-percent level, respectively.

explanatory variable—distance—on the conditional probability of becoming listed publicly on the stock market. A positive estimated coefficient would indicate a higher value of the hazard rate and therefore a positive impact on the likelihood of the firm going public at that point in time. Because the geographic measure reflects firm proximity to an university, we expect a negative sign, which would indicate that firms locating closer to an university should endure a shorter duration between formation and the IPO. According to Table 6-5, the negative coefficient on the distance between the firm and the nearest university confirms Hypothesis 6.

In addition, the duration to IPO is lower when the firm is located close to universities with a high number of graduates in the natural sciences. This would again suggest that accessing the tacit knowledge embodied in human capital requires geographic proximity in the natural sciences but not in the social sciences. Also, as mentioned by McWilliams and Siegel (2000) and Audretsch and Lehmann (2005b), the investments in R&D improve performance, in this study ensured by the time-to-IPO.

## 6. CONCLUSIONS

A recent literature has emerged suggesting that not only are the spillovers of knowledge important in generating innovative output but that universities provide an important source of such knowledge spillovers (Hall, Link and Scott, 2003). However, this literature has generally ignored the impact that such university spillovers exert in shaping the strategic location decisions of firms. The results of this study not only confirm that university spillovers play an important role, but also that they have a strong influence in the strategic location decisions of firms. In particular, the empirical evidence suggests that geographic proximity is a key element of firm strategy. However, the location decision is shaped not only by the output of universities, but also by the nature of that output. In this paper, we consider two specific university outputs—research and education, which generates human capital—in two different fields, the natural sciences and the social sciences. To access knowledge transmitted by published articles in the natural sciences, geographic proximity is not particularly important. This is consistent with the findings of Audretsch and Stephan (1996) that geographic proximity is not a prerequisite to access and absorb codified knowledge. By contrast, in the social sciences, geographic proximity to the universities is apparently more important, which may reflect a higher tacit knowledge content in social science research that reflects the lack of a unified scientific methodology.

These results are actually reversed in accessing the educational output of universities, in the form of graduated students. Firms tend to locate in

geographic proximity to universities with a high number of graduates in the natural sciences, which presumably indicates the limited geographic options for students with human capital specific to particular technologies. By contrast, firms are less geographically restricted with respect to universities with a high number of graduates in the social sciences, which reflects the more general skills and human capital rendering the student more mobile.

This paper also provides at least some evidence that the strategic location choice shapes firm performance. In particular, we find that the duration between start-up and going public is less when the firm is located within geographic proximity to the university.

These results confirm the resource-based view of the firm in that research and human capital are important resources shaping not just the location decision but also the performance of new and young firms. These resources can be obtained through accessing spillovers of knowledge from sources that are external to the firm, in this case from the university. However, as the results of this paper suggest, the strategic location decision of the firm and the role of geographic proximity will be shaped not only by the existence of knowledge spillovers but also the particular type of knowledge spillover. The location decision to access knowledge in the natural sciences clearly has different strategic implications than to absorb knowledge emanating from the social sciences. In both cases, geography plays a role, albeit a decidedly different one. However, the results are based on small and medium sized and publicly-held firms and thus may not be generalizable for other firms, like privately held firms or countries.

## NOTES

<sup>1</sup> See also Stephan et al. (2002), analyzing the new venture's placement of PhD students.

<sup>2</sup> However, this may yield to an over-estimate of the impact of knowledge spillovers and/or the overall importance of geographical proximity to this group of new ventures.

<sup>3</sup> The publications in social science and natural science did not vary across the universities during time.

<sup>4</sup> However, the results did not vary according the assumption of the underlying distribution of the variables.

<sup>5</sup> The University of Ulm (University of Erfurt) has no students in social science (natural sciences).

## REFERENCES

- Acs, Z. J., Audretsch, D. B. and Feldman M. P. (1991). Real Effects of Academic Research: Comment. *American Economic Review* 82: 363–367.
- Acs, Z. J., Audretsch, D. B. and Feldman, M. P. (1994). The Recipients of R&D Spillovers: Firm Size and Innovation. *Review of Economics and Statistics* 76: 336–340.

- Agarwal, R., Echambadi, R. and Sarkar, M. B. (2002). The Conditioning Effect of Time on Firms Survival: A Life Cycle Approach. *Academy of Management Journal* 45 (8): 971–994.
- Alvarez, S. A. and Barney, J. B. (2003). Organizing Rent Generation and Appropriation: Toward a Theory of the Entrepreneurial Firm, *Journal of Business Venturing* 19: 612–635.
- Anselin, L., Varga, A. and Acs, Z. (1997). Local Geographic Spillovers between University Research and High technology Innovations. *Journal of Urban Economics* 42: 422–448.
- Anselin, L., Varga, A. and Acs, Z. (2000). Geographical Spillovers and University Research: A Spatial Econometric Perspective. *Growth and Change* 31: 501–515.
- Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity*, R. Nelson (Ed.). Princeton University Press, Princeton.
- Audretsch, D. B. (2007). *The Entrepreneurial Society*. Oxford University Press, New York.
- Audretsch, D. B. and Lehmann, E. E. (2005a). Does the New Economy Need New Governance? Incentives, Knowledge, and Performance. *Review of Accounting and Finance* 4: 13–33.
- Audretsch, D. B. and Lehmann, E. E. (2005b). Do Spillovers Pay? *Economics of Innovation and New Technology* 15: 71–81.
- Audretsch, D. B. and Lehmann, E. E. (2004). Debt or Equity: The Role of Venture Capital in Financing the New Economy in Germany. *Schmalenbach Business Review* 56: 340–357.
- Audretsch, D. B., Lehmann, E. E. and Warning, S. (2005a). University Spillover and New Firm Location. *Research Policy* 34: 1113–1122.
- Audretsch, D. B., Lehmann, E. E. and Warning, S. (2005b). University Spillovers: Does the Kind of Science Matter? *Industry and Innovation* 11 (3): 193–206.
- Audretsch, D. B. and Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86: 630–640.
- Audretsch, D. B. and Stephan, P. E. (1999). Knowledge Spillovers in Biotechnology: Sources and Incentives. *Journal of Evolutionary Economics* 19: 97–107.
- Audretsch, D. B., Keilbach, M. and Lehmann, E. E. (2006). *Entrepreneurship and Economic Growth*. Oxford University Press, New York.
- Bania, N., Eberts, R. and Fogerty, M. S. (1993). Universities and the Start-up of New Companies: Can we Generalize from Route 128 and Silicon Valley? *Review of Economics and Statistics* 75: 761–766.
- Bartik, T. J. (1985). Business Location Decisions in the United States: Estimation of the Effects of Unionization, Taxes, and Other Characteristics of States, *Journal of Business and Economic Statistics* 3: 14–22.
- Baum, J. A. C. and Sorenson, O. (2003). *Advances in Strategic Management: Geography and Strategy*, Vol. 20, JAI Press, Greenwich, CT (forthcoming).
- Becattini, G. (1990). The Marshallian Industrial District as a Socio-Economic Notion. In *Industrial Districts and Inter-Firm Co-operation in Italy*, F. Pyke, G. Becattini and W. Sengenberger (Eds), pp. 37–51. International Institute for Labour Studies, ILO, Geneva.
- Brusco, S. (1990). The Idea of the Industrial District. Its Genesis In *Industrial Districts and Inter-Firm Co-operation in Italy*, F. Pyke, G. Becattini and W. Sengenberger (Eds), pp. 10–19. International Institute for Labour Studies, ILO, Geneva.
- Carlton, D. (1983). The Location and Employment Choices of New Firms. *Review of Economics and Statistics* 65: 440–449.
- Chevassus-Lozza, E. and Galliano, D. (2000). *Firm's Location and Competitiveness: The Case of the French Food Industry*, Paper presented at the 6<sup>th</sup> World Congress of Regional Science Association International, Lugano.
- Cohen, W. M. and Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* 35: 128–152.
- Dumais, G., Ellisson, G. and Glasear, E. L. (2002). Geographic Concentration as a Dynamic Process. *Review of Economics and Statistics* 84: 193–204.



- Evans, D. S. and Leighton, L. S. (1989). Some Empirical Aspects of Entrepreneurship. *American Economic Review* 79: 519–535.
- Fabel, O. (2004). The Emergence of a New Economy: An O-Ring Approach. *Journal of Institutional and Theoretical Economics* (forthcoming).
- Feldman, M. (1994). Knowledge Complementary and Innovation. *Small Business Economics* 6: 363–372.
- Feldman, M. (1994). *The Geography of Innovation*. Kluwer, Boston.
- Feldman, M. P. (2000). Location and Innovation: The New Economic Geography of Innovation. In *Handbook of Economic Geography*, G. Clark, M. P. Feldman and M. Gertler (Eds). Oxford University Press, Oxford.
- Gartner, W. and Carter, N. M. (2003). Entrepreneurial Behavior and Firm Organizing Processes. In *International Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds). Kluwer Academic, Dordrecht.
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy* 100: 1126–1152.
- Grabher, G. (1993). The Weakness of Strong Ties: The Lock-In of Regional Developments in the Ruhr Area. In *The Embedded Firm—On the Socioeconomics of Industrial Networks*, G. Grabher (Ed.), pp. 255–277. Routledge, London.
- Greene, W. H. (2003). *Econometric Analysis*, 5<sup>th</sup> edition. Prentice Hall, New York.
- Griliches, Z. (1979). Issues in Assessing the Contribution of R&D to Productivity Growth. *Bell Journal of Economics* 10: 92–116.
- Hall, B. H., Link, A. N. and Scott, J. T. (2003). Universities as Research Partners. *Review of Economics and Statistics* 85: 485–491.
- Harhoff, D. (1999). Firm Formation and Regional Spillovers—Evidence from Germany. *Economics of Innovation & New Technology* 8 (1/2): 27–56.
- Harhoff, D. (2000). R&D Spillovers, Technological Proximity, and Productivity Growth—Evidence from German Panel Data. *Schmalenbach Business Review* 52: 238–260.
- Harris, R. G. (2001). The Knowledge-Based Economy: Intellectual Origins and New Economic Perspectives. *International Journal of Management Review* 3 (1): 21–41.
- Henderson, R. and Cockburn, I. (1994). Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal* 15: 63–84.
- Henderson, R., Jaffe, A. and Trajtenberg, M. (1998). Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting 1965–1988, *Review of Economics and Statistics* 65: 119–127.
- Hoang, H. and Antoncic, B. (2003). Network-Based Research in Entrepreneurship: A Critical Review. *Journal of Business Venturing* 18: 165–187.
- Holmes, T. J. and Schmitz, J. A. (1990). A Theory of Entrepreneurship and its Application to the Study of Business Transfers. *Journal of Political Economy* 98: 265–294.
- Jaffe, A. B. (1989). Real Effects of Academic Research. *American Economic Review* 79: 957–970.
- Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as evidenced by Patent Citations. *Quarterly Journal of Economics* 63: 577–598.
- Jovanovic, B. (1994). Entrepreneurial Choice When People Differ in their Management and Labor Skills. *Small Business Economics* 6: 185–192.
- Kihlstrom, R. and Laffont, J.-J. (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy* 87: 719–748.
- Kiefer, N. (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature* 26: 646–679.
- Kim, P. H., Aldrich, H. E. and Keister, L. A. (2003). *If I Were Rich? The Impact of Financial and Human Capital on Becoming a Nascent Entrepreneur*. Discussion Paper, Dept. of Sociology, University of North Carolina at Chapel Hill.

- Knight, F. (1921). *Risk, Uncertainty and Profit*. Augustus Kelly, New York.
- Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organizational Science* 3: 383–397.
- Krueger, N. F., Jr. (2003). Cognitive Psychology of Entrepreneurship. In *International Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds). Kluwer Academic, Dordrecht.
- Krugman, P. (1991). *Geography and Trade*. MIT Press, Cambridge, MA.
- Lazear, E. P. (2002). *Entrepreneurship*. Working Paper No. 9109. National Bureau of Economic Research, Cambridge, MA.
- Liebeskind, J., Oliver, A. L., Zucker, L. G. and Brewer, M. (1996). Social Networks, Learning, and Flexibility: Sourcing Scientific Knowledge in New Biotechnology Firms. *Organizational Science* 7: 428–443.
- Link, A. N. and Scott, J. C. (2005). The Growth of Research Triangle Parks. *Small Business Economics* (forthcoming).
- Lucas, R. E. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics* 9: 508–523.
- Lucas, R. E. (1993). Making a Miracle. *Econometrica* 61: 251–272.
- MacPherson, A. D. (1998). Academic-Industry Linkages and Small Firm Innovation: Evidence from the Scientific Instruments Sector. *Entrepreneurship and Regional Development* 10: 261–276.
- Marshall, A. (1890). *Principles of Economics*. MacMillan, London.
- McWilliams, A. and Siegel, D. S. (2002). Corporate Social Responsibility and Financial Performance: Correlation or Misspecification? *Strategic Management Journal* 21: 603–609.
- Mowery, D. C. and Ziedonis, A. A. (2001). *The Geographic Reach of Market and Non-Market Channels of Technology Transfer: Comparing Citations and Licenses of University Patents*. Working paper No. 8568. National Bureau of Economic Research, Washington, DC.
- Parker, S. C. (2003). Asymmetric Information, Occupational Choice and Government Policy. *Economic Journal* 113: 861–882.
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Piore, M. J. and Sabel, C. F. (1984). *The Second Industrial Divide. Possibilities for Prosperity*. Basic Books, New York.
- Porter, M. E. (2000). Locations, Clusters, and Company Strategy. In *The Oxford Handbook of Economic Geography*, G. L. Clark, M. P. Feldman and M. S. Gertler (Eds), pp. 253–274. Oxford University Press, Oxford.
- Porter, M. E. (1990). *The Comparative Advantage of Nations*. Free Press, New York.
- Pyke, F., Becattini, G. and Sengenberger, W. (1990). *Industrial Districts and Inter-Firm Cooperation in Italy*. International Institute for Labor Studies, ILO, Geneva.
- Reynolds, P. D., Storey, D. J. and Westhead, P. (1994). Cross National Comparison of the Variation on the New Firm Formation Rates. *Regional Studies* 27: 443–456.
- Santoro, M. D. and Chakrabarti, A. K. (2002). Firm Size and Technology Centrality in Industry-University Interactions. *Research Policy* 31: 1163–1180.
- Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Rte. 128*. Harvard University Press, Cambridge, MA.
- Schrader, S. (1991). Informal Technology Transfer between Firms: Cooperation through Information Trading. *Research Policy* 20: 153–170.
- Shane, S. (2000). Prior Knowledge and the Discovery of Entrepreneurial Opportunities. *Organization Science* 11: 448–469.
- Shane, S. (2001a). Technological Opportunities and New Firm Creation. *Management Science* 47: 205–220.

- Shane, S. (2001b). Technology Regimes and New Firm Formation. *Management Science* 47: 1173–1190.
- Shane, S. and Venkataraman, S. (2000). The Promise of Entrepreneurship As a Field of Research. *Academy of Management Review* 26: 216–226.
- Shane, S. and Eckhardt, J. (2003). The Individual-Opportunity Nexus. In *Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds), pp. 161–194. Kluwer, Boston.
- Siegel, D. S., Westhead, P. and Wright, M. (2003). Assessing the Impact of Science Parks on the Research Productivity of Firms: Exploratory Evidence from the United Kingdom. *International Journal of Industrial Organization* 21: 1357–1369.
- Sorensen, O. and Audia, G. (2000). The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the U.S. 1940–1989. *American Journal of Sociology* 106: 324–362.
- Spence, M. A. (1984). Cost Reduction, Competition, and Industry Performance. *Econometrica* 52: 101–121.
- Stephan, P. (1996). The Economics of Science. *Journal of Economic Literature* 34: 1199–1235.
- Stephan, P. E., Sumell, A. J., Black, G. C. and Adams, J. D. (2002). *Public Knowledge, Private Placements: New Ph.D.s as a Source of Knowledge Spillovers*. Working Paper, Georgia State University.
- Stevenson, H. H. and Jarillo, J. C. (1990). A Paradigm of Entrepreneurship: Entrepreneurial Management. *Strategic Management Journal* 11: 17–27.
- Stuart, T. E. and Sorenson, O. (2003). The Geography of Opportunity: Spatial Heterogeneity in Founding Rates and the Performance of Biotechnology Firms. *Research Policy* (forthcoming).
- Stuart, T. E. and Shane, S. (2002). Organizational Endowments and the Performance of University Start-ups. *Management Science* 48: 151–170.
- Stuart, T. E., Hoang, H. A. and Hybels, R. C. (1999). Inter-organizational Endorsements and the Performance of Entrepreneurial Ventures. *Administrative Science Quarterly* 44: 315–349.
- Thornton, P. H. and Flynn, K. H. (2003). Entrepreneurship, Networks and Geographies. In *International Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds). Kluwer Academic, Dordrecht.
- Varga, A. (2000). Local Academic Knowledge Transfers and the Concentration of Economic Activity. *Journal of Regional Science* 40: 289–309.
- Von Hippel, E. (1994). Sticky Information and the Locus of Problem Solving: Implications for Innovation. *Management Science* 40: 429–439.
- Warning, S. (2004). Performance Differences in German Higher Education. Empirical Analysis of Strategic Groups. *Review of Industrial Organization* 24: 393–408.
- Zucker, L. G., Darby, M. R. and Armstrong, J. (1998). Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises. *American Economic Review* 88: 290–306.

MARTIN A. CARREE

*Department of Organization and Strategy, University of Maastricht, FdEWB, PO Box 616, 6200 MD Maastricht, The Netherlands*

*E-mail: m.carree@os.unimaas.nl*

## 7. On Factors Promoting and Hindering Entry and Exit

### 1. INTRODUCTION

The importance of new entry for market competition, efficiency and economic development is largely undisputed. Alfred Marshall (1961) used the parable of the young trees of the forest replacing large old trees that gradually lose their vitality. The deteriorating economic performance of centrally planned economies showed the consequences of frustrating entrepreneurial initiatives. The lack of new entry may also have been a problem for the Japanese economy in the past decade (see Kawai and Urata, 2002). New entrants can bring innovative business solutions to the market, sometimes even leading to the foundation of completely new industries (for a recent example of a software industry, see Giarratana, 2004). They may notice profit opportunities that are overlooked by incumbent firms and increase market efficiency. Industries with low birth and death rates are likely to be more vulnerable to an inadequate allocation of resources, limited innovativeness and some form of formal or tacit collusion (Geroski and Jacquemin, 1985). Therefore, high barriers to entry and exit may be serious impediments to dynamic market efficiency.

The aim of this chapter is to discuss and evaluate the empirical evidence on the processes of entry and exit that has accumulated since the early contribution of Mansfield (1962). There is relatively little empirical work on entry and exit compared to the huge amount of theoretical work in the Industrial Organization literature (Disney et al., 2003). Firm entry and exit rates vary widely across *industries* (industrial dimension); see, for example, Dunne et al. (1988). There are more barriers to starting up or closing down an airplane manufacturing company than a restaurant. But barriers also differ strongly over *time* (temporal dimension) within an industry following the industry life cycle;

see, for example, Gort and Klepper (1982). And entry and exit rates (aggregated or within an industry) may differ from one *region* (regional dimension) to another; see, for example, Audretsch and Fritsch (1999) and Carree (2002). As a specific case of the latter category, there may be differences in entry and exit rates between *countries*, for example, due to cultural factors; see, for example, Mueller and Thomas (2000) and Reynolds et al. (2002). Finally, the probability of entry or exit can differ across *individuals* (individual dimension) within the same region. Persons with high financial, human and social capital may be more likely to start a venture and make it successful compared to persons who lack such resources. This chapter seeks to provide an overview of the factors behind these differences in entry and exit rates and on their interrelationship. Specific attention will be paid to the factors promoting, and the factors hindering, entry and exit in regions and industries.

## 2. ENTREPRENEURSHIP AND ENTRY

New venture creation is traditionally regarded as being at the heart of the research field of entrepreneurship (Gartner, 1985, 1990; Low and McMillan, 1988). However, more recently, studies of entrepreneurial behavior have been extended to include corporate entrepreneurship (or intrapreneurship). Lumpkin and Dess (1996) argued that the essential act of entrepreneurship is new entry. New entry, in their opinion, is defined as entering new or established markets with new or existing products. Hence, this may be achieved by starting a business, but also through an existing business (intrapreneurship). Nevertheless, founding a firm is widely regarded as a prime example of entrepreneurial activity (Verheul et al., 2005). The extent of new venture creation differs vastly across countries. This can be derived from data collected through the Global Entrepreneurship Monitor (GEM); see, for example, Reynolds et al. (2002).

The Adult Population Survey of the GEM measures the *total entrepreneurial activity* rate (TEA), defined as the percentage of adult population (18–64 years old) that is either actively involved in starting a new venture or the owner/manager of a business that is less than 42 months old. This percentage ranges from close to 20% for Thailand and India to less than 3% for Japan and Russia (see Table 7-1). There are various reasons for the differences in entrepreneurial activity rates across countries. In developing countries micro-enterprises (in the informal sector) can be set up and dissolved with very limited means. There is a lot of entrepreneurial activity in these countries, but the vast majority of these “enterprises” remains very small. In some former communist countries, like Poland, Slovenia, Croatia and Russia, entrepreneurial activity may be low since the population did not grow up in a society in which entrepreneurship played a role.

TABLE 7-1 *Total entrepreneurial activity rates (TEA) across countries, 2002*

| Country     | TEA   | Country      | TEA   | Country   | TEA   |
|-------------|-------|--------------|-------|-----------|-------|
| Thailand    | 0.189 | Norway       | 0.087 | Poland    | 0.044 |
| India       | 0.179 | Israel       | 0.071 | Taiwan    | 0.043 |
| Chile       | 0.157 | Switzerland  | 0.071 | Sweden    | 0.040 |
| Korea       | 0.145 | Hungary      | 0.066 | Croatia   | 0.036 |
| Argentina   | 0.142 | Denmark      | 0.065 | Hong Kong | 0.034 |
| New Zealand | 0.140 | South Africa | 0.065 | France    | 0.032 |
| Brazil      | 0.135 | Italy        | 0.059 | Belgium   | 0.030 |
| Mexico      | 0.124 | Singapore    | 0.059 | Russia    | 0.025 |
| China       | 0.123 | UK           | 0.054 | Japan     | 0.018 |
| Iceland     | 0.113 | Germany      | 0.052 |           |       |
| U.S.        | 0.105 | Finland      | 0.046 |           |       |
| Ireland     | 0.091 | Netherlands  | 0.046 |           |       |
| Canada      | 0.088 | Slovenia     | 0.046 |           |       |
| Australia   | 0.087 | Spain        | 0.046 |           |       |

Source: GEM.

Entrants are usually small (see, e.g., Geroski, 1995). They perform an essentially entrepreneurial task. They may see previously unnoticed profit opportunities and try to capitalize on this knowledge by starting a venture. Kirzner (1973, 1997) stressed this role. They may be innovative in terms of their product, organization of production and combination of resources. Hence, they may be the prime cause of economic development as discussed by Schumpeter (1934). The reason why the vast majority of ventures start small-scale has been argued to be self-selection in the initial commitments by entrepreneurs (see Caves, 1998). A real option perspective suggests that whereas entrepreneurs may start out small when they expect their chances of success to be low, at the same time, small-scale entry commonly provides an option to invest heavily if early returns are promising. Entrants holding more positive expectations about their capabilities are likely to make larger initial commitments.

### 2.1. *Who Enters? The Individual Dimension*

An important element connecting entrepreneurship to entry is the question: Who enters? What are the characteristics of entrepreneurs who start new ventures? Can we predict whether a certain individual would be more likely to become an entrepreneur than another individual? Four main factors have been considered in the literature. These are psychological factors and human, social and financial capital. Some aspects of each of those four categories will be discussed below.<sup>1</sup>

Three important psychological factors that have been connected to entrepreneurship are need for achievement, locus of control and risk aversion. McClelland (1961) introduced the notion of the need for achievement as a key characteristic of successful entrepreneurs. Individuals with a stronger desire to strive for excellence are assumed to be more likely to become entrepreneurs. Rotter (1966) presented another psychological trait: locus of control. People attribute the reason for their performance either to themselves or to external factors. Those who assume it to be largely dependent upon their own actions have an internal locus of control. They are assumed to be more likely to start a venture than individuals with a more external locus of control. In a recent longitudinal study, Hansemark (2003) found no evidence that a need for achievement affected the probability of new start-ups, but did find evidence that a locus of control had predictive power (but only for men). Individuals who are more averse to risk are assumed to be less likely to start up an enterprise with its inherent uncertainties. Khilstrom and Laffont (1979) derived, in a neoclassical framework, how the least risk-averse individuals become entrepreneurs. However, the model does not allow for individuals to also become part-time self-employed, considerably reducing the risk of variation in income over time. The empirical results with respect to risk aversion are unclear and mixed (Parker, 2004, pp. 83–84). Other psychological factors mentioned in the literature include love for autonomy and personal perseverance.

Human capital is the collection of personal abilities and knowledge. It is usually measured through (years of) education and (years of) experience. The direction of the effect of both education and experience on the probability of becoming self-employed is not entirely obvious. The same skills that would make a person a good entrepreneur may also make him an employee with a very promising career in a large corporation or government institution. Therefore, the effect of education on self-employment is likely to be industry dependent. Bates (1995), for example, found positive effects for services, but negative effects for construction. Parker (2004, p. 73) reported mixed results in the empirical research into the effect of education on the probability of self-employment. With regard to experience it is important to discriminate between paid-employment experience and self-employment experience. Evans and Leighton (1989) reported that previous self-employment experience has a positive effect on the probability of entering self-employment, with previous paid-employment experience having no effect. Davidsson and Honig (2003) found empirical support for each of the effects of education, work experience and start-up experience, with the latter having the strongest effect (see also Shane, 2001). Lazear (2002, 2004) recently suggested that entrepreneurs are jacks-of-all-trades, not excelling in any one skill but competent in many. Entrepreneurs must have sufficient knowledge in a variety of areas to survive and be successful, while employees can usually specialize much more in the

specific job they take. Wagner's (2003) empirical results support Lazear's claims.

An extensive social network can also be considered a form of capital: "social capital." High social capital provides entrepreneurs with access to information and cooperation and trust from others. Baron and Markman (2003) distinguished social capital from social competence, which is the ability to interact effectively with others. Social capital is far from evenly distributed in society. Anderson and Miller (2003) discussed how entrepreneurs from higher socio-economic classes enjoy enhanced access to effective business support and to opportunities. Davidsson and Honig (2003) found that having parents in business or close friends or neighbors in business has a substantial positive effect on the probability of being a nascent entrepreneur. The effect of having parents in business may of course affect the decision to become self-employed in different ways, for example, by inheriting the business or by the skills learned while working in the family company as a youngster. Davidsson and Honig also found that only one aspect of social capital, namely, being a member of a business network, affects outcomes including the first sale or showing a profit. Brüderl and Preisendörfer (1998) showed that social network support is positively related to survival and profitability of recently started ventures. Carree and Verheul (2005) found that entrepreneurs with more entrepreneurial contacts devote more hours to their company.

The fourth form of capital is financial capital. In their influential study, Evans and Jovanovic (1989) found empirical evidence for binding liquidity constraints: many individuals are prevented from trying entrepreneurship because of lack of access to financial resources. This is confirmed by a series of papers, for example, Holtz-Eakin et al. (1994a) and Van Praag and Van Ophem (1995). Personal wealth may not only increase the probability of entry into self-employment but can also lower the probability of exit. Holtz-Eakin et al. (1994b) found inheritances to increase the probability of survival. See Parker (2004, chapter 7) for an extensive overview of the empirical evidence on credit rationing.

## 2.2. *Who Exits?*

The same individual-level factors that influence entry are likely to influence survival or exit. For example, previous self-employment experience is found to positively affect the probability of survival (e.g., Holmes and Schmitz, 1996). Also, access to capital is found to positively affect business survival (e.g., Bates, 1990). In addition to such factors, two basic characteristics of the firm are widely confirmed to have a positive effect on staying in business: its age and its size (see Parker, 2004, pp. 222–223). There is both a liability of newness and of smallness. This is in line with the predictions made by the



Jovanovic (1982) passive learning model. Storey and Wynarczyk (1996) find that firm characteristics—age, size, sector, location—are more important than human capital for explaining survival.

### 3. INCENTIVES FOR ENTRY AND EXIT

Entry and exit rates differ widely across industries, over time and across regions. In Table 7-2 the entry and exit rates (both in terms of number of establishments and employment) are shown for U.S. industries. The entry and exit rates for U.S. industries, in terms of number of establishments, average 11.5% and 10.8%, respectively. The employment impact of entering and exiting firms is only about half of that. Some sectors show much more entry (e.g., information) than others (e.g., manufacturing).<sup>2</sup> There are two key reasons underlying these patterns: difference in incentives to enter and differences in the barriers to enter. Why it can be more attractive to enter one industry or

TABLE 7-2 *Entry and exit rates for U.S. industries (2nd digit NAICS)*

| Industry (NAICS)              | Establishment |              | Employment    |              | Establishment<br>number<br>(1000s) | Employment<br>number<br>(1000s) |        |
|-------------------------------|---------------|--------------|---------------|--------------|------------------------------------|---------------------------------|--------|
|                               | Entry<br>rate | Exit<br>rate | Entry<br>rate | Exit<br>rate |                                    |                                 |        |
| Total                         | 0.115         | 0.108        | 0.059         | 0.053        | 6297                               | 114,034                         |        |
| Agriculture etc.              | 11            | 0.170        | 0.139         | 0.102        | 0.088                              | 22                              | 184    |
| Mining                        | 21            | 0.110        | 0.095         | 0.055        | 0.042                              | 21                              | 456    |
| Utilities                     | 22            | 0.095        | 0.081         | 0.055        | 0.035                              | 17                              | 655    |
| Construction                  | 23            | 0.125        | 0.124         | 0.055        | 0.060                              | 608                             | 6572   |
| Manufacturing                 | 31            | 0.076        | 0.087         | 0.023        | 0.034                              | 332                             | 16,475 |
| Wholesale trade               | 42            | 0.090        | 0.102         | 0.052        | 0.057                              | 410                             | 6112   |
| Retail trade                  | 44            | 0.098        | 0.097         | 0.058        | 0.048                              | 1024                            | 14,843 |
| Transport & warehousing       | 48            | 0.146        | 0.139         | 0.048        | 0.049                              | 167                             | 3791   |
| Information                   | 51            | 0.198        | 0.147         | 0.105        | 0.077                              | 118                             | 3546   |
| Finance & insurance           | 52            | 0.132        | 0.132         | 0.100        | 0.075                              | 392                             | 5965   |
| Real estate & rental & leas.  | 53            | 0.128        | 0.109         | 0.092        | 0.070                              | 264                             | 1944   |
| Prof., scient. & techn. serv. | 54            | 0.140        | 0.119         | 0.078        | 0.065                              | 620                             | 6819   |
| Manag. of comp. & enterp.     | 55            | 0.115        | 0.113         | 0.063        | 0.063                              | 45                              | 2874   |
| Administrative/support serv.  | 56            | 0.147        | 0.130         | 0.088        | 0.086                              | 304                             | 9139   |
| Educational services          | 61            | 0.119        | 0.088         | 0.021        | 0.018                              | 61                              | 2534   |
| Health care & social assis.   | 62            | 0.094        | 0.076         | 0.040        | 0.034                              | 610                             | 14,111 |
| Arts, entertainment & recr.   | 71            | 0.123        | 0.108         | 0.054        | 0.046                              | 85                              | 1742   |
| Accommodation & foodserv.     | 72            | 0.119        | 0.112         | 0.080        | 0.063                              | 483                             | 9880   |
| Other services                | 81            | 0.083        | 0.083         | 0.043        | 0.042                              | 665                             | 5296   |

Note: source is U.S. Small Business Administration, for 2000–2001. Left out categories are 95 (auxiliaries) with 14,363 establishments and 99 (unclassified) with 35366 establishments, respectively. Entry and exit rates are measured in terms of establishments and employment.

TABLE 7-3 Summary of empirical studies into entry and exit determinants

|             | Effect of profitability |   |    |    |    | Effect of growth |   |   |    |    |
|-------------|-------------------------|---|----|----|----|------------------|---|---|----|----|
|             | --                      | - | 0  | +  | ++ | --               | - | 0 | +  | ++ |
| Gross entry | 1                       | 0 | 14 | 11 | 9  | 0                | 0 | 9 | 10 | 18 |
| Gross exit  | 1                       | 3 | 3  | 2  | 1  | 4                | 1 | 4 | 0  | 0  |
| Net entry   | 0                       | 1 | 4  | 6  | 2  | 0                | 1 | 4 | 1  | 12 |

Source: Carree and Thurik (1996), ++ means *t*-value higher than +2.5, + means *t*-value between +1.5 and +2.5, 0 means *t*-value between -1.5 and +1.5, - means *t*-value between -2.5 and -1.5 and -- means *t*-value less than -2.5.

region versus another is dealt with in the current section. Barriers are discussed in the next section.

Entrepreneurs or (diversifying) firms enter an industry or region if they perceive that they are better off than by refraining from entry. The most obvious incentive would be that entry is profitable. Two obvious candidates for variables that could capture the extent of *ex-post* profitability (*ex-post* meaning after entry has taken place) are *ex-ante* profitability and the growth rate of demand. Highly profitable, strongly growing markets usually appeal to new entrants. The empirical evidence for the effect of market growth is relatively strong. In a meta-study, Carree and Thurik (1996) found that, out of 37 empirical studies incorporating the effect of a measure of *growth* on *gross* entry, no less than 28 report a positive and significant effect (none report a significant negative effect). In addition, in 13 out of 18 studies investigating the impact on *net* entry, a positive and significant effect was found (one finds a significant negative effect). The evidence for a positive effect of current profitability is only slightly weaker. Carree and Thurik found 20 out of 35 empirical studies to have presented evidence for a significant positive effect of a measure of *profitability* on *gross* entry (and only one a significant negative effect). In addition, in eight out of 13 studies investigating the impact on *net* entry, profitability has a significant and positive effect (and again only one a significant negative effect). See Table 7-3.

### 3.1. The Impact of Current Profitability

Profitability is undoubtedly an attraction to potential entrants. However, the extent to which *ex-ante* profitability is always an adequate measure of *ex-post* profit is unclear. Kessides (1990), for example, noted that the defense of high rents may lead incumbent firms to threaten post-entry retaliation. There are other reasons for *ex-ante* profitability to have limited impact on entry (and exit). Entrepreneurs may not be that aware of profit opportunities available and, hence, there may be limited competition in the sense of “the free entry of rivals, each in an incessant race to better the others” (Ikeda, 1990, p. 79). The amount

of entrepreneurial activity may be just too low to quickly adjust profits to long-run equilibrium levels. For example, Geroski and Masson (1987) estimated the speed of the competitive process of excess profits disappearing over time to be very slow. Geroski (1995) claimed that a slow reaction of entry to high profits is a stylized fact in the empirical literature on entry.

The positive effect of profitability on *gross* entry as depicted in Table 7-3 holds across data for different countries and time periods. Examples include the first study by Mansfield (1962) for four U.S. industries over the 1916–59 period, Highfield and Smiley (1987) for U.S. data (60 industries) in the 1976–81 period, Schwalbach (1987) for diversifying entry in Germany in the 1977–82 period, Khemani and Shapiro (1988) for Canadian data in the 1972–76 period, Rosenbaum and Lamort (1992) for U.S. data (213 industries) in the 1972–82 period, Carree and Thurik (1996) for Dutch retail industries in the 1981–88 period and Amel and Liang (1997) for U.S. local banking markets in the period 1977–88. However, there are also empirical studies that were unable to find a positive and significant effect. Examples include Hamilton (1985) for Scottish industries in the 1976–80 period, Mata (1993) for Portuguese data of both specialist and diversifying entry in the 1982–86 period and Santarelli and Sterlacchini (1994) for Italian industries in the 1986–89 period. Some of the disparity in findings may be due to some entry barriers not being incorporated into the analysis. High profits may not attract any entry when barriers are high.

### 3.2. *The Impact of Market Growth*

Growing markets are attractive for entrants because these markets are usually characterized by less vigorous competition and by many emerging market niches. Incumbent firms sometimes cannot keep track with increasing demand, leaving market room to entrants. Hause and Du Rietz (1984) built on this notion by suggesting a nonlinear (convex) effect of market growth: a doubling of the growth rate leads, *ceteris paribus*, to a more than doubling of the entry rate. Empirical studies that find *no* effect of market growth on *gross* entry are the exception: see Table 7-3. Examples include Masson and Shaanan (1982) for U.S. industries in the 1958–63 period, Baldwin and Gorecki (1987) for foreign entry in Canadian manufacturing in the 1970s and Audretsch and Acs (1994) for U.S. industries in the 1976–86 period. Jackson (1984) provides support for the notion that a growing market is likely to lead to more market niches. He shows how the number of commodities expands with total expenditure both in the aggregate and for commodity groups.

The size of the effect of market growth on entry is not easily comparable across studies. However, Carree and Thurik (1999) found a similar demand elasticity for *net* entry in Dutch retailing in the 1980s of about 0.3 as found earlier by Acs and Audretsch (1989) and Hirschey (1981) using comparable measures

for U.S. manufacturing. However, such demand elasticity is somewhat of an oversimplification since the composition of demand may also be important. In addition it matters a lot whether there is a certain growth rate of demand in the very early stages of the industry life cycle or later on. We turn to this issue below.

### 3.3. The Impact of the Industry Life Cycle

The literature on industry life cycles is very important for the question of when entries and exits occur in time. Important contributions to this literature include Gort and Klepper (1982), Klepper and Graddy (1990), Agarwal and Gort (1996) and Klepper (1996). The U.S. tire industry has been especially focused upon because of the wealth of data available for this industry; see Carree and Thurik (2000), Jovanovic and MacDonald (1994) and Klepper and Simons (2000). The industry life cycle has five separate stages; see, for example, Agarwal and Gort (1996). Figure 7-1 presents these five stages. The first stage is one of introduction, there is only entry, at a relatively moderate pace and there is virtually no exit. The entry rate peaks in the second stage and the exit rate slowly starts to rise. During the third stage entry, which is on the decline, and exit, which is on the increase, are about equal. The exit rate peaks in the fourth stage of shakeout. And in the final, fifth, stage there is still some entry and exit, but mainly in niches and the industry life cycle has reached full maturity.

Klepper (1996, p. 562) summarized the life cycle as follows: “When industries are new, there is a lot of entry, firms offer many different versions of

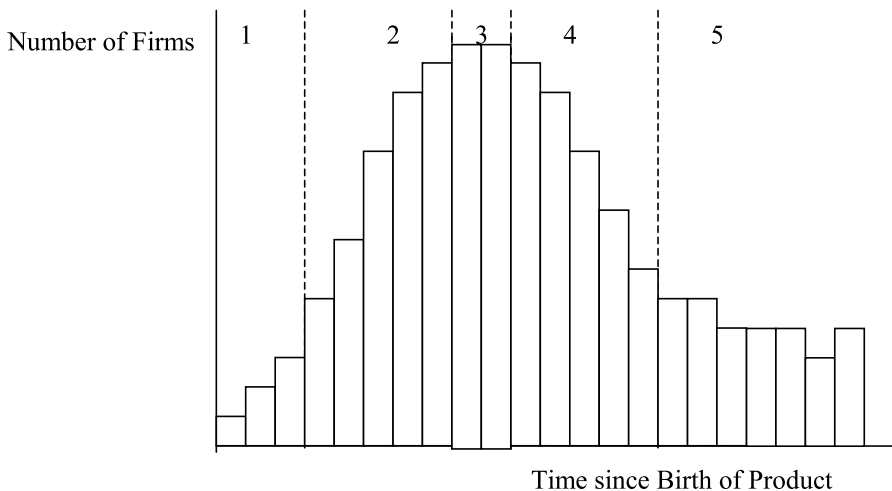


FIGURE 7-1 The life cycle of an industry: five stages.

the industry's product, the rate of product innovation is high, and market shares change rapidly. Despite continued market growth, subsequently entry slows, exit overtakes entry and there is a shakeout in the number of producers, the rate of product innovation and the diversity of competing versions of the product decline, increasing effort is devoted to improving the production process, and market shares stabilize." The industry life cycle theory suggests *a negative correlation between entry and exit rates over time*. This correlation is the consequence of a structural process often taking several decades. The industry life cycle can also be used to indicate what type of firms enter and exit; see, for example, Karlsson and Nystrom (2003) investigating the knowledge-intensity of firms for Swedish manufacturing data.

Klepper (1996) argued that there is an important first-mover advantage for early entrants. In his eyes, early entrants are among the most likely to dominate the industry later on. Geroski (2003) mentioned four different types of first-mover advantages. The first is the head start in traveling along learning curves and exploiting economies of scale. The second is that first movers may have the opportunity to monopolize scarce inputs. The third is a consumer lock-in effect: consumers may be reluctant to change products they have grown accustomed to. The fourth is the enhanced brand identity and status resulting from being the first to the market. There are, of course, also second-mover advantages: learning from the mistakes of predecessors and free riding on their efforts. Even so, the advantages for early entrants make it more difficult to achieve successful entry later on during the industry life cycle.

### *3.4. The Impact of Technological Opportunities*

Shane (2001) stressed the importance of technological opportunities on firm formation. He investigated empirically the impact of the importance, radicalness and patent scope of an invention on the likelihood of the formation of a new firm. Prusa and Schmitz (1991) had already suggested that new firms might be better at radical innovation than incumbents. Shane confirmed this finding in his study. In addition, there are the findings by Acs and Audretsch (1988) that when an industry is composed of large firms, innovative activity will tend to emanate more from the small firms than from the large firms. Hence, innovation appears to be a viable gateway to entry. Audretsch (1995) provided empirical evidence that small firms' innovation rate positively affects the rate of entry into industries. Industries in their early stages of the life cycle with many technological opportunities are likely to attract many new entrants. However, as Caves (1998, p. 1969) noted: "the opportunity to make and appropriate

innovations, a gateway to entry early in the process, becomes a barrier to entry in the mature stage.”

### 3.5. Regions: Agglomeration or Competition

The presence of many incumbents in a certain region is likely to affect the incentives for entry and exit. However, the effect on incentives may differ from one industry to another. In some industries, large numbers of incumbents promote entry and discourage exit because these regions display *agglomeration* (positive clustering) effects. In other industries large number of incumbents will discourage new entry and encourage exit owing to pressures of *competition*. An example of an industry in which agglomeration effects are important is hi-tech. An example of an industry in which the competition effect is likely to dominate is consumer services.

Belderbos and Carree (2002), studying the determinants of location of Japanese investments in China, and Barry et al. (2003), studying those of location of U.S. firms in Ireland, both claimed that there are substantial agglomeration effects. The probability of Japanese electronics firms of investing in Chinese regions was found by Belderbos and Carree to be positively affected by the existence of (Japanese) electronics firms. Barry et al. distinguished between efficiency effects of agglomeration and demonstration effects whereby existing firms send signals to new investors as to the attractiveness of the region and found empirical evidence for the existence of both sources of agglomerations. Beaudry (2001) provided empirical evidence of strong positive clustering effects in the U.K. aerospace industry leading to new entry. Acs and Armington (2004) found a relationship between the local levels of human capital and firm formation rates, stressing the importance of human capital externalities.

However, the reverse effect of regions with many firms being unattractive for entry has also been found. See, for example, Acs and Audretsch (1989) for evidence in U.S. manufacturing industries that small firms do not tend to enter industries in which there is already a considerable presence of small firms. Both Carree and Thurik (1999) and Carree (2002) applied an error correction framework predicting the extent of net entry in retail and consumer service markets and confirms that industries and regions with many firms are less attractive for entry. This is in line with the survey by Geroski (1995) concluding that net entry should be represented as an error-correction process that renders further entry unprofitable. The sheer number of firms already in the market increases the risk that a new entrant will not succeed. Hence, the importance of the finding by Fan and White (2003) that regions with low bankruptcy exemptions levels receive more entry: in highly competitive markets, entry will only take place when exit barriers are relatively low.

### 3.6. Regions: The Impact of Unemployment

As noted above, there is a range of individual-level factors affecting the decision to enter self-employment. A personal situation that may lead to firm formation is unemployment. Evans and Leighton (1990) showed evidence that unemployed workers are about twice as likely to start businesses as employed workers. This may be reason to suspect that regions with high unemployment rates are also characterized by high entry (and low exit) rates. However, evidence for this is mixed at best. Storey (1991) suggested that, in general, time-series analyses point to unemployment being positively related to indices of new firm formation, while cross-sectional studies indicate the opposite (see also Foti and Vivarelli, 1994, p. 83). Audretsch and Fritsch (1999) find some evidence, for a dataset of German regions, of unemployment push effects for small-scale easy-to-enter industries. Carree (2002) found little evidence for the unemployment push hypothesis with the possible exception of a couple of very easy-to-enter industries like used merchandise stores and automotive repair shops. An important reason for the lack of evidence for the unemployment push hypothesis is that, first, unemployment may be an indicator of a depressed economic environment which cannot be completely controlled for by adding business cycle variables; and second, the unemployed may have less human (or entrepreneurial) capital on average when compared to the employed. For example, Acs and Armington (2004) reported a positive impact of higher local proportions of adults with college degrees on rates of new firm formation.

## 4. BARRIERS TO ENTRY AND EXIT

Entrants have to surmount barriers to entry, while firms that wish to exit may have to deal with barriers to exit. Industries that are characterized by high barriers to entry usually also have high barriers to exit. Entrepreneurs who make large investments entering an industry will not be tempted to leave the industry early, risking a lot of investment (sunk cost) lost. This is one of the reasons why *entry and exit rates are positively correlated cross-sectionally*. New firms (greenfield entrants) will be especially scared off by barriers to entry. This may be less so for diversifying entrants. R&D and advertising may generate externalities that can be used efficiently in adjacent industries inducing firms to enter similar industries (see, e.g., Sembenelli and Vannoni, 2000).

### 4.1. Entry Barriers

Some industries, like musical instrument stores, beauty shops and automotive repair shops are relatively easy to enter. These industries are often characterized by an absence of important scale economies, by limited start-up

capital and by the absence of technological complexities. Other industries, like car manufacturing, manufacturing of microprocessors and nuclear plants, have much higher barriers to entry. Firms in these industries deal with complex hi-tech production processes that cannot be easily copied by potential entrants to the industry. The barriers discussed above are called *structural* (exogenous) barriers. They are barriers that are not erected by incumbent firms but result from the specific product or production process in the industry at hand. There are also *strategic* (endogenous) barriers: barriers deliberately erected by market participants to forestall entry. An example is patents.

Entry barriers are not constant over time. New barriers may arise while others disappear over the life cycle of an industry. In the early stages of the life cycle of an industry, entry barriers may be low. See, for example, Klepper (2002) for the case of the car manufacturing industry. Over time, barriers like technological complexity, consumer loyalty and economies of scale in production and R&D are likely to increase. There are few examples of structural barriers that tend to decline over time in an industry. An example may be that of the emergence of a dominant design (Utterback and Abernathy, 1975; Suarez and Utterback, 1995) which slows down technological advances and thereby decreases the barrier of uncertainty for new entrants about which kind of product they should choose to produce.

Entry barriers may also differ from one (local) region or country to another. There may be regions that subsidize new entry, for example, by opening up new business parks with favorable conditions for firms. Countries may differ widely in their entry regulations. Fonseca et al. (2001) showed that the number of procedures for a start-up and the average time until start-up differs widely across countries, with Germany, Italy and Spain, for instance, having substantial start-up costs, while countries like Denmark, United Kingdom and United States impose low start-up costs. A lowering of *legal* barriers in a country may provoke additional entry, especially so in the short term. Deregulation in U.S. and European airlines has led to new entry. Ingham and Thompson (1995) showed how deregulation in financial services has created a spurt of entry. Carree and Nijkamp (2001) showed that the removal of institutional barriers to entry in the Netherlands led to increased entry in retail industries. Fan and White (2003) found that the probability of households owning businesses is 35% higher if they live in U.S. states with unlimited rather than low bankruptcy exemptions levels. The interpretation is that higher exemption levels benefit potential entrepreneurs who are risk averse.

Many different forms of entry barriers have been discussed in the literature. Shepherd (1997, p. 210) lists 22 different types of entry barriers, while Karakaya and Stahl (1989) provide a survey of 19 different market entry barriers. The three most well-known types of entry barriers were introduced by



Bain (1956). There is much less literature available on exit barriers. For a paper specifically focusing upon these barriers, see Karakaya (2000).

#### 4.2. Bainian Barriers to Entry

Bain (1956) provided a seminal analysis of barriers to entry. He considered systematically potential competition in addition to competition from existing rivals. Bain considered entry barriers to be anything that allows incumbents to earn above-normal profits without inducing entry. He distinguished three categories of entry barriers: absolute cost advantages, product differentiation and scale economies. These three different types of barriers are discussed below. Specific attention is paid to the issue of first-mover advantage.

There can be several reasons why incumbents may have *absolute cost advantages* over potential entrants. Reasons for a cost advantage include learning by doing and the results of R&D. When innovations are protected by patents, new entrants are denied access to the superior production process or product. Cost advantages may also result from incumbents being able to buy inputs, including investment capital, at lower prices than entrants. When absolute cost advantages exist, the entrant faces higher costs than the incumbent, with the latter being able to make a profit. First movers in the market obviously can benefit from learning by doing and by filing the first patent applications. They may also secure access to strategic inputs.

Firms seek to avoid price competition by differentiating their products. Bain stresses that advertising is an important means of *product differentiation*, especially in consumer goods industries. Advertising increases customer loyalty making it harder for a new entrant to gain market share. Schmalensee (1982) claimed that pioneering firms gain familiarity among customers who are then reluctant to switch. Szymanski et al. (1995) performed a meta-analysis and find that, on average, earlier entry is associated with greater market share.

The presence of substantial *scale economies* requires entrants to produce at a substantial scale immediately upon entering a market. When economies of scale are sizable, entering below the minimum efficient scale will lead to higher unit costs than the large incumbent firms. In some industries entering at for example half the minimum efficient scale leads to considerably higher costs. Shepherd (1997) mentioned examples like synthetic rubber, commercial aircraft and electric motors. Klepper (1996) argued that there are scale economies in R&D leading to large enterprises being able to produce superior products or to produce at lower cost than smaller counterparts, which leads, in the long run, to a “shake-out” of small firms. First movers have the advantage of increasing the scale of production (and R&D) when potential entrants are not even present in the market.

### 4.3. *Barriers or Pathways?*

It has been argued by some (e.g., Caves, 1998) that there may be a “barrier-versus-gateway duality.” This means that some variables may under some circumstances function as a gateway to entry and under other circumstances as a barrier. The example of innovation being a gateway to entry in industries that are either young or already dominated by large firms has already been given above. Another example is advertising. Advertising has been, in general, considered as a barrier to entry. The accumulation of advertising leads to a goodwill entry barrier in many consumer goods industries. However, Kessides (1986) claimed that when demand is apparently considerably affected by advertising, this may provide an opportunity for new entrants financially capable of advertising to gain market share. Yet another example is patents. When strong patent protection is possible this can be a strong barrier to entry since imitation may be blocked. On the other hand, for new ventures such patent protection may be vital (Shane, 2001).

### 4.4. *Strategic Entry Barriers*

Incumbent firms may limit price, install excess capacity or have numerous patents to forestall entry. These barriers to entry are called strategic barriers since they are the deliberate choice of incumbent firms. Bunch and Smiley (1992) performed questionnaire research on nine different possible strategies and found evidence of the use of strategic entry deterrents, especially in concentrated and R&D-intensive industries. They also found that firms expend fewer resources on entry deterrence when other barriers to entry are present. The most common strategic barriers were found to be the creation of product loyalty through advertising, filling product niches, masking the results for highly profitable divisions and patent preemption. Capacity preemption and limit pricing are least often used. See also Smiley (1988). Lieberman (1987) also found that incumbents rarely build excess capacity preemptively in an effort to deter entry. His sample was one of U.S. chemical product industries. Little empirical evidence for the presence of limit pricing is available and it is heavily criticized from a theoretical point of view by claiming that it is an irrational strategy (Lipczynski and Wilson, 2001). Chang and Tang (2001) confirmed for Singapore that strategies of advertising, filling product niches and hiding profits, next to dominating distribution channels, are often used to blockade entry. Thomas (1999) showed that advertising is used in the ready-to-eat cereal industry to limit the scale of entry. Thomas also shows that entrants are likely to be met with an aggressive price reaction.

Dixit (1982) developed a simple game-theoretic model showing that the profitability of entry will be affected by the incumbent’s subsequent actions.

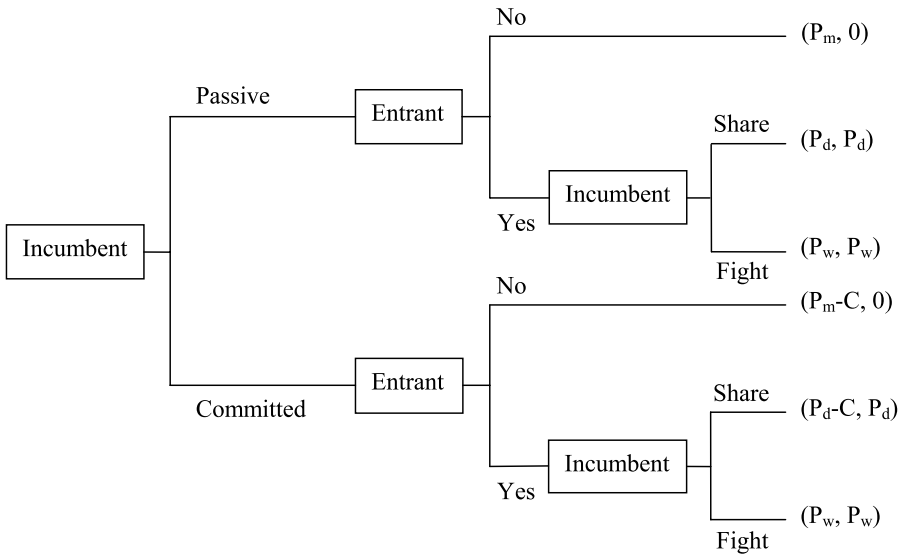


FIGURE 7-2 *The Dixit (1982) game tree.*

The two-stage game tree is shown in Figure 7-2: the first profit between brackets in the figure is for the incumbent; the second is for the entrant. There are two parties involved, an incumbent monopolist and a potential entrant. Let us start with the top part of the game tree. When the potential entrant stays out, the monopolist earns a profit  $P_m$ . However, when there is entry, the two parties may share the market (duopoly) and earn a profit  $P_d$ . The incumbent may also choose to fight a price war, which is mutually destructive, both earning  $P_w$ . It is assumed that  $P_m > P_d > 0 > P_w$ . In the top part of the game tree (the “passive incumbent”) the outcome would be that there is entry since the entrants knows that the incumbent’s optimal response to entry is sharing (and  $P_d > 0$ ). Assume now that the incumbent is not “passive” but has available a prior irrevocable commitment, like excess capacity, which incurs a cost  $C$  in readiness to fight a price war. When a price war occurs, however, this cost does not emerge (since for example capacity is completely used). A “committed” incumbent finds it optimal to fight the price war in the event of entry when  $P_w > P_d - C$ . Hence, the potential entrant, aware of this, will stay out since a price war results in  $P_w$ , which is negative. The incumbent, knowing this in turn, will choose to make the commitment when the outcome of the bottom part of the game tree exceeds that of the top part:  $P_m - C > P_d$ . So, under the condition that there is a cost  $C$  for which  $P_m - P_d > C > P_d - P_w$  the incumbent will strategically erect an entry barrier. It has a credible threat to potential entrants into the market. The social cost of this strategic barrier is substantial: there is the lack of competition (loss of consumers’ surplus) and there is the excess capacity (resource cost

of the commitment instrument). The assumptions of the above example also suggest why in practice excess capacity is rarely built deliberately. An important assumption is that there is one incumbent, whereas in reality in most industries there are many. The question is then, given firms acting independently, which of the firms will erect this excess capacity.

## 5. MODELING ENTRY AND EXIT AND THEIR INTERRELATIONSHIP

The empirical modeling of entry (and exit) took off with the work of Orr (1974) (see also Shapiro and Khemani, 1987). The basic model developed is that entry (or exit) is a function of (i) barriers to entry (exit), (ii) current opportunities and (iii) controls. The current opportunities are usually measured by profitability and market growth. The equation for the number of entrants (exiting firms) is:

$$\text{Entry (or Exit)} = F(\text{Barriers, Current Opportunities, Controls}).$$

This relation stresses the key point that there has to be both willingness and opportunity to enter.<sup>3</sup> One of the controls is usually the size of the market (e.g., number of incumbents). An example of a simple model in this context would be:

$$\text{Entry}_t / \text{Incumbents}_{t-1} = (a + b \times \text{Barrier}_t) \times \text{Profit}_{t-1}.$$

This equation relates the entry rate in period  $t$  to the profit rate in the previous period. The extent to which the entry rate reacts to this profit rate is assumed to be dependent upon the height of the entry barrier. In the absence of barriers, one would expect entrepreneurs to quickly react to profitability:  $a > 0$ . In the presence of barriers, the speed of reaction to profits will be lower or there be no reaction at all if barriers are insurmountable:  $b < 0$ .

This equation has several disadvantages. A first important disadvantage is that strategic (endogenous) entry barriers cannot be incorporated simply into this model. Such entry barriers function *ex ante* via the threat of post-entry incumbent reprisals. A second important disadvantage is that the dynamic interaction between entry and exit is not taken into account.

Births and deaths may be interrelated not only because of the underlying industry life cycle process, or because of barriers in the market, but also because one causes the other. When entry causes exit, this is called *displacement*. When exit causes entry, this is called *replacement*. The effect of entry (exit) in one period leading to entry (exit) in a consecutive period is called a *demonstration* effect. A range of papers has investigated the dynamic and/or simultaneous interrelationship between entry and exit. Examples include Rosenbaum and

Lamort (1992) using U.S. data, Johnson and Parker (1994) using U.K. data, Carree and Thurik (1996) using data on Dutch retailing, Kangasharju and Moisio (1998) for Finnish regions, Fotopoulos and Spence (1998) using Greek manufacturing data and Lay (2003) using data on Taiwanese manufacturing.

Replacement, displacement and demonstration effects (of a possibly complex inter-temporal nature) can be incorporated into the model by having:

$$\begin{aligned} \text{Entry}_t &= F(\text{Entry}_{t-1} \dots \text{Entry}_{t-T}, \text{Exit}_t \dots \text{Exit}_{t-T}, \text{Barriers}, \\ &\quad \text{Current Opportunities, Controls}), \\ \text{Exit}_t &= G(\text{Exit}_{t-1} \dots \text{Exit}_{t-T}, \text{Entry}_t \dots \text{Entry}_{t-T}, \text{Barriers}, \\ &\quad \text{Current Opportunities, Controls}). \end{aligned}$$

The estimation of a simultaneous relationship between entry and exit is a complicated venture. Rosenbaum and Lamort (1992), Carree and Thurik (1996) and Fotopoulos and Spence (1998) all end up facing the same dilemma: although the system should in principle be estimated with a simultaneous equations estimator like 3SLS, the hypothesis of no simultaneity cannot be rejected, so that an estimation technique like SUR can be used. Another problem is that in many cases entry and exit are intimately connected. For example, in some countries a firm sold from one owner to the next is recorded as an exit and an entry, while in others it is not. A third problem is that of multicollinearity among the entry and exit variables and between these variables and the barriers, opportunities and controls. A final problem is that barriers are often (relatively) constant over time leading to similar estimation problems as in dynamic panel data models with dummies. These kind of methodological and data problems make comparison across studies difficult, although, in general, support for displacement and replacement effects has been claimed.

## 6. CONCLUSION

Entry and exit rates can differ widely between industries, between regions, between individuals and over time. They are key features of the dynamics of industries and regions. There are various reasons for these differences. The current chapter relates entry (and exit) to entrepreneurship and discusses a range of factors that make one industry (or region) more likely to attract entrants or to have more firms exiting than others. These factors are related to incentives, barriers and the dynamic interrelationship between entry and exit. The chapter covers general patterns: it must be acknowledged of course that there is a huge difference between a small part-time retail venture opening up and a diversified entry employing hundreds of workers. The entry barriers they face are very

different and so are the opportunities they exploit. Still, the chapter provides a general overview of factors behind the processes of entry and exit, at the industrial, temporal, regional and individual level.

## NOTES

<sup>1</sup> Also the start-up size of firms may be influenced by individual-level factors, see e.g. Colombo et al. (2004).

<sup>2</sup> Dunne et al. (1988) report average annual entry and exit rates for U.S. manufacturing for the 1963–1982 period of 8.1% and 7.4%, respectively, close to the figures in Table 7-2 for manufacturing.

<sup>3</sup> See Van Praag and Van Ophem (1995) for a study that discriminates between willingness and opportunity in case of self-employment. They find that there are almost seven times more individuals who wish to switch to self-employment than the actual number of switchers.

## REFERENCES

- Acs, Z. J. and Armington, C. (2004). The Impact of Geographic Differences in Human Capital on Service Firm Formation Rates. *Journal of Urban Economics* 56: 244–278.
- Acs, Z. J. and Audretsch, D. B. (1988). Innovation in Large and Small Firms: An Empirical Analysis. *American Economic Review* 78: 678–690.
- Acs, Z. J. and Audretsch, D. B. (1989). Small-Firm Entry in US Manufacturing. *Economica* 56: 255–265.
- Agarwal, R. and Gort, M. (1996). The Evolution of Markets and Entry, Exit and Survival of Firms. *Review of Economics and Statistics* 78: 489–498.
- Amel, D. F. and Liang, J. N. (1997). Determinants of Entry and Profits in Local Banking Markets. *Review of Industrial Organization* 12: 59–78.
- Anderson, A. R. and Miller, C. J. (2003). “Class matters”: Human and Social Capital in the Entrepreneurial Process. *Journal of Socio-Economics* 32: 17–36.
- Audretsch, D. B. (1995). *Innovation and Industry Evolution*. MIT Press, Cambridge, MA.
- Audretsch, D. B. and Acs, Z. J. (1994). New-Firm Startups, Technology, and Macroeconomic Fluctuations. *Small Business Economics* 6: 439–449.
- Audretsch, D. B. and Fritsch, M. (1999). The Industry Component of Regional New Firm Formation Processes. *Review of Industrial Organization* 15: 239–252.
- Bain, J. S. (1956). *Barriers to New Competition*. Harvard University Press, Cambridge, MA.
- Baldwin, J. R. and Gorecki, P. K. (1987). Plant Creation versus Plant Acquisition. *International Journal of Industrial Organization* 5: 27–41.
- Baron, R. A. and Markman, G. D. (2003). Beyond Social Capital: The Role of Entrepreneurs’ Social Competence in Their Financial Success. *Journal of Business Venturing* 18: 41–60.
- Barry, F., Gorg, H. and Strobl, E. (2003). Foreign Direct Investment, Agglomerations, and Demonstration Effects: An Empirical Investigation. *Weltwirtschaftliches Archiv* 139: 583–600.
- Bates, T. (1990). Entrepreneur Human Capital Inputs and Small Business Longevity. *Review of Economics and Statistics* 72: 551–559.
- Bates, T. (1995). Self-employment Entry across Industry Groups. *Journal of Business Venturing* 10: 143–156.

- Beaudry, C. (2001). Entry, Growth and Patenting in Industrial Clusters: A Study of the Aerospace Industry in the UK. *International Journal of the Economics of Business* 8: 405–436.
- Belderbos, R. and Carree, M. (2002). The Location of Japanese Investments in China: Agglomeration Effects, Keiretsu, and Firm Heterogeneity. *Journal of the Japanese and International Economies* 16: 194–211.
- Brüderl, J. and Preisendörfer, P. (1998). Network Support and the Success of Newly Founded Businesses. *Small Business Economics* 10: 213–225.
- Bunch, D. S. and Smiley, R. (1992). Who Deters Entry? Evidence on the Use of Strategic Entry Deterrents. *Review of Economics and Statistics* 74: 509–521.
- Carree, M. A. (2002). Does Unemployment Affect the Number of Establishments? A Regional Analysis for US States. *Regional Studies* 36: 389–398.
- Carree, M. A. and Nijkamp, J. (2001). Deregulation in Retailing: the Dutch Experience. *Journal of Economics and Business* 53: 225–235.
- Carree, M. and Thurik, R. (1996). Entry and Exit in Retailing: Incentives, Barriers, Displacement and Replacement. *Review of Industrial Organization* 11: 155–172.
- Carree, M. A. and Thurik, A. R. (1999). The Carrying Capacity and Entry and Exit Flows in Retailing. *International Journal of Industrial Organization* 17: 985–1007.
- Carree, M. A. and Thurik, A. R. (2000). The Life Cycle of the U.S. Tire Industry. *Southern Economic Journal* 67: 254–278.
- Carree, M. and Verheul, I. (2005). Time Allocation by the Self-Employed: The Determinants of the Number of Working Hours in Start-Ups. *Applied Economics Letters*, forthcoming.
- Caves, R. E. (1998). Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature* 36: 1947–1982.
- Chang, H. L. and Tang, F. F. (2001). An Empirical Study on Strategic Entry Barriers in Singapore. *Asia Pacific Journal of Management* 18: 503–517.
- Colombo, M. G., Delmastro, M. and Grilli, L. (2004). Entrepreneurs' Human Capital and the Start-Up Size of New Technology-Based Firms. *International Journal of Industrial Organization* 22: 1183–1211.
- Davidsson, P. and Honig, B. (2003). The Role of Social and Human Capital among Nascent Entrepreneurs. *Journal of Business Venturing* 18: 301–331.
- Disney, R., Haskel, J. and Heden, Y. (2003). Entry, Exit and Establishment Survival in UK Manufacturing. *Journal of Industrial Economics* 51: 91–112.
- Dixit, A. (1982). Recent Developments in Oligopoly Theory. *American Economic Review Papers and Proceedings* 72: 12–17.
- Dunne, T., Roberts, M. J. and Samuelson, L. (1988). Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *RAND Journal of Economics* 19: 495–515.
- Evans, D. S. and Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *Journal of Political Economy* 97: 808–827.
- Evans, D. S. and Leighton, L. S. (1989). Some Empirical Aspects of Entrepreneurship. *American Economic Review* 79: 519–535.
- Evans, D. S. and Leighton, L. S. (1990). Small Business Formation by Unemployed and Employed Workers. *Small Business Economics* 2: 319–330.
- Fan, W. and White, M. J. (2003). Personal Bankruptcy and the Level of Entrepreneurial Activity. *Journal of Law and Economics* 46: 543–567.
- Fonseca, R., Lopez-Garcia, P. and Pissarides, C. A. (2001). Entrepreneurship, Start-up Costs and Employment. *European Economic Review* 45: 692–705.
- Foti, A. and Vivarelli, M. (1994). An Econometric Test of the Self-Employment Model: The Case of Italy. *Small Business Economics* 6: 81–93.
- Fotopoulos, G. and Spence, N. (1998). Entry and Exit From Manufacturing Industries: Symmetry, Turbulence and Simultaneity—Some Empirical Evidence from Greek Manufacturing Industries, 1982–1988. *Applied Economics* 30: 245–262.

- Gartner, W. B. (1985). A Conceptual Framework for Describing the Phenomenon of New Venture Creation. *Academy of Management Review* 10: 696–706.
- Gartner, W. B. (1990). What Are We Talking About When We Talk About Entrepreneurship? *Journal of Business Venturing* 5: 15–28.
- Geroski, P. A. (1995). What Do We Know About Entry? *International Journal of Industrial Organization* 13: 421–440.
- Geroski, P. A. (2003). *The Evolution of New Markets*. Oxford University Press, Oxford.
- Geroski, P. A. and Jacquemin, A. (1985). Industrial Change, Barriers to Mobility, and European Industrial Policy. *Economic Policy* 1: 169–204.
- Geroski, P. A. and Masson, R. T. (1987). Dynamic Market Models in Industrial Organization. *International Journal of Industrial Organization* 5: 1–13.
- Giarratana, M. S. (2004). The Birth of a New Industry: Entry by Start-Ups and the Drivers of Firm Growth: The Case of Encryption Software. *Research Policy* 33: 787–806.
- Gort, M. and Klepper, S. (1982). Time Paths in the Diffusion of Product Innovations. *Economic Journal* 92: 630–653.
- Hamilton, R. T. (1985). Interindustry Variation in Gross Entry Rates of ‘Independent’ and ‘Dependent’ Business. *Applied Economics* 17: 271–280.
- Hansemark, O. C. (2003). Need for Achievement, Locus of Control and the Prediction of Business Start-Ups: A Longitudinal Study. *Journal of Economic Psychology* 24: 301–319.
- Hause and Du Rietz (1984). Entry, Industry Growth, and the Micro-Dynamics of Industry Supply. *Journal of Political Economy* 92: 733–757.
- Highfield, R. and Smiley, R. (1987). New Business Starts and Economic Activity: An Empirical Investigation. *International Journal of Industrial Organization* 5: 51–66.
- Hirschey, M. (1981). The Effect of Advertising on Industrial Mobility, 1947–72. *Journal of Business* 54: 329–339.
- Holmes, T. J. and Schmitz, J. A. (1996). Managerial Tenure, Business Age, and Small Business Turnover. *Journal of Labor Economics* 14: 79–99.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. S. (1994a). Entrepreneurial Decisions and Liquidity Constraints. *RAND Journal of Economics* 25: 334–347.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. S. (1994b). Sticking it Out: Entrepreneurial Survival and Liquidity Constraints. *Journal of Political Economy* 102: 53–75.
- Ikeda, S. (1990). Market-Process Theory and ‘Dynamic’ Theories of the Market. *Southern Economic Journal* 57: 75–92.
- Ingham, H. and Thompson, S. (1995). Deregulation and Diversifying Entry Decisions: The Case of Financial Services. *Review of Economics and Statistics* 77: 177–183.
- Jackson, L. F. (1984). Hierarchic Demand and the Engle Curve for Variety. *Review of Economics and Statistics* 66: 8–15.
- Johnson, P. and Parker, S. (1994). The Interrelationships Between Births and Deaths. *Small Business Economics* 6: 283–290.
- Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica* 50: 649–670.
- Jovanovic, B. and MacDonald, G. M. (1994). The Life Cycle of a Competitive Industry. *Journal of Political Economy* 102: 322–347.
- Kangasharju, A. and Moisio, A. (1998). Births-Deaths Nexus of Firms: Estimating VAR with Panel Data. *Small Business Economics* 11: 303–313.
- Karakaya, F. (2000). Market Exit and Barriers to Exit: Theory and Practice. *Psychology & Marketing* 17: 651–668.
- Karakaya, F. and Stahl, M. J. (1989). Barriers to Entry and Market Entry Decisions in Consumer and Industrial Goods Markets. *Journal of Marketing* 53: 80–91.
- Karlsson, C. and Nystrom, K. (2003). Exit and Entry over the Product Life Cycle: Evidence from the Swedish Manufacturing Industry. *Small Business Economics* 21: 135–144.



- Kawai, H. and Urata, S. (2002). Entry of Small and Medium Enterprises and Economic Dynamism in Japan. *Small Business Economics* 18: 41–51.
- Kessides, I. N. (1986). Advertising, Sunk Costs, and Barriers to Entry. *Review of Economics and Statistics* 68: 84–95.
- Kessides, I. N. (1990). Towards a Testable Model of Entry: A Study of the US Manufacturing Industries. *Economica* 57: 219–238.
- Khemani, R. S. and Shapiro, D. M. (1988). On Entry and Mobility Barriers. *Antitrust Bulletin* 33: 115–134.
- Khilstrom, R. E. and Laffont, J. J. (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy* 87: 719–748.
- Kirzner, I. M. (1973). *Competition & Entrepreneurship*. University of Chicago Press, Chicago.
- Kirzner, I. M. (1997). Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature* 35: 60–85.
- Klepper, S. (1996). Entry, Exit, Growth, and Innovation over the Product Life Cycle. *American Economic Review* 86: 562–583.
- Klepper, S. (2002). The Capabilities of New Firms and the Evolution of the US Automobile Industry. *Industrial and Corporate Change* 11: 645–666.
- Klepper, S. and Graddy, E. (1990). The Evolution of New Industries and the Determinants of Market Structure. *RAND Journal of Economics* 21: 27–44.
- Klepper, S. and Simons, K. L. (2000). The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the U.S. Tire Industry. *Journal of Political Economy* 108: 728–760.
- Lay, T. J. (2003). The Determinants of and Interaction Between Entry and Exit in Taiwan's Manufacturing. *Small Business Economics* 20: 319–334.
- Lazear, E. P. (2002). *Entrepreneurship*. NBER Working Paper 9109, National Bureau of Economic Research.
- Lazear, E. P. (2004). Balanced Skills and Entrepreneurship. *American Economic Review Papers and Proceedings* 94: 208–211.
- Lieberman, M. B. (1987). Excess Capacity as a Barrier to Entry: An Empirical Appraisal. *Journal of Industrial Economics* 35: 607–627.
- Lipczynski, J. and Wilson, J. (2001). *Industrial Organization: An Analysis of Competitive Markets*. Prentice Hall, Harlow.
- Low, M. and MacMillan, I. C. (1988). Entrepreneurship: Past Research and Future Challenges. *Journal of Management* 14: 139–161.
- Lumpkin, G. T. and Dess, G. G. (1996). Clarifying the Entrepreneurial Orientation Construct and Linking it to Performance. *Academy of Management Review* 21: 135–172.
- Mansfield, E. (1962). Entry, Gibrat's Law, Innovation, and the Growth of Firms. *American Economic Review* 52: 1023–1051.
- Marshall, A. (1961). *Principles of Economics*. MacMillan, London.
- Masson, R. T. and Shaanan, J. (1982). Stochastic-Dynamic Limit Pricing: An Empirical Test. *Review of Economics and Statistics* 64: 413–422.
- Mata, J. (1993). Entry and Type of Entrant: Evidence from Portugal. *International Journal of Industrial Organization* 11: 101–122.
- McClelland, D. C. (1961). *The Achieving Society*. Van Nostrand, Princeton.
- Mueller, S. L. and Thomas, A. S. (2000). Culture and Entrepreneurial Potential: A Nine Country Study of Locus of Control and Innovativeness. *Journal of Business Venturing* 16: 51–75.
- Orr, D. (1974). The Determinants of Entry: A Study of the Canadian Manufacturing Industries. *Review of Economics and Statistics* 56: 58–66.
- Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*. Cambridge University Press, Cambridge.

- Praag, C. M. van and Van Ophem, H. (1995). Determinants of Willingness and Opportunity to Start as an Entrepreneur. *Kyklos* 48: 513–540.
- Prusa, T. J. and Schmitz, J. A., Jr. (1991). Are New Firms an Important Source of Innovation? Evidence from the PC Software Industry. *Economics Letters* 35: 339–342.
- Reynolds, P. D., Bygrave, W. D., Autio, E., Cox, L. W. and Hay, M. (2002). *Global Entrepreneurship Monitor 2002 Executive Report*. Babson College, Wellesley, MA.
- Rosenbaum, D. I. and Lamort, F. (1992). Entry, Barriers, Exit, and Sunk Costs: An Analysis. *Applied Economics* 24: 297–304.
- Rotter, J. B. (1966). Generalized Expectancies for Internal versus External Control of Reinforcement. *Psychological Monographs* 80: 1–28.
- Santarelli, E. and Sterlacchini, A. (1994). New Firm Formation in Italian Industry: 1985–89. *Small Business Economics* 6: 95–106.
- Schmalensee, R. (1982). Product Differentiation: Advantages of Pioneering Brands. *American Economic Review* 72: 349–365.
- Schumpeter, J. A. (1934). *The Theory of Economic Development*. Harvard University Press, Cambridge, MA.
- Schwalbach, J. (1987). Entry by Diversified Firms into German Industries. *International Journal of Industrial Organization* 5: 43–49.
- Sembenelli, A. and Vannoni, D. (2000). Why Do Established Firms Enter Some Industries and Exit Others? Empirical Evidence on Italian Business Groups. *Review of Industrial Organization* 17: 441–456.
- Shane, S. (2001). Technological Opportunities and New Firm Creation. *Management Science* 47: 205–220.
- Shapiro, D. and Khemani, R. S. (1987). The Determinants of Entry and Exit Reconsidered. *International Journal of Industrial Organization* 5: 15–26.
- Shepherd, W. G. (1997). *The Economics of Industrial Organization*, 4<sup>th</sup> edition. Prentice Hall, London.
- Smiley, R. H. (1988). Empirical Evidence on Strategic Entry Deterrence. *International Journal of Industrial Organization* 6: 167–180.
- Storey, D. J. (1991). The Birth of New Firms—Does Unemployment Matter? A Review of the Evidence. *Small Business Economics* 3: 167–178.
- Storey, D. J. and Wynarczyk, P. (1996). The Survival and Non Survival of Micro Firms in the UK. *Review of Industrial Organization* 11: 211–229.
- Suarez, F. F. and Utterback, J. M. (1995). Dominant Designs and the Survival of Firms. *Strategic Management Journal* 16: 415–430.
- Szymanski, D. M., Troy, L. C. and Bharadwaj, S. G. (1995). Order of Entry and Business Performance: An Empirical Synthesis and Reexamination. *Journal of Marketing* 59: 17–33.
- Thomas, L. A. (1999). Incumbent Firms' Response to Entry: Price, Advertising, and New Product Introduction. *International Journal of Industrial Organization* 17: 527–555.
- Utterback, J. M. and Abernathy, W. J. (1975). A Dynamic Model of Process and Product Innovation. *Omega* 3: 639–656.
- Verheul, I., Uhlaner, L. and Thurik, R. (2005). Business Accomplishments, Gender and Entrepreneurial Self-Image. *Journal of Business Venturing* 20: 483–518.
- Wagner, J. (2003). Testing Lazear's Jack-of-all-Trades View of Entrepreneurship with German Micro Data. *Applied Economics Letters* 10: 687–689.

## Stage 3: Financing Ventures

ROBERT CRESSY

*Cass Business School, London EC1Y 8TZ, United Kingdom*

*E-mail: r.c.cressy@city.ac.uk*

## 8. Debt Finance and Credit Constraints on SMEs

### 1. INTRODUCTION

It is well established in the academic literature that bank debt is the most important source of external finance for small firms (see, e.g., Cressy, 1993, 1996; Petersen and Rajan, 1994; Cosh and Hughes, 1994; Berger and Udell, 1998). These studies detail the characteristics of, and the role played by, debt in small firm finance and operations. We shall in this chapter draw on these and other sources in characterizing loan finance for small businesses. However, for some years there has been a widespread and growing belief both in academia and in policy circles that small firms do not get *enough* bank funding. Indeed, by far the *majority* of the theoretical literature on small firm borrowing is focused on this question, following some seminal papers in the early 1980s, Stiglitz and Weiss (1981) being perhaps the most celebrated. These considerations make an exploration of lending constraints on small firms central to an examination of the role of bank debt in small business finance. Despite the emphasis on credit constraints, in this chapter we shall not, however, attempt survey in any detail the vast theoretical literature on the subject as there are already several good surveys available (see Parker, 2002, 2004). Our task here will instead be confined to surveying *testable* theories or, more precisely, *tested* theories, and specifically those with a focus on bank lending constraints on small businesses.<sup>1</sup>

The remainder of the chapter is structured as follows. We begin with some definitions of key terms that will be used in the survey. For purposes of exposition we also divide the small firm population broadly into two categories, the *typical* and the *sophisticated* small firm (and entrepreneur), respectively. In Section 2 we identify the main sources of small firm finance, internal and external to the firm. In Sections 3 and 4 we define loan collateral and loan term features more generally, outlining the rationale for banks demanding collateral

for loans to small businesses. In Section 5 we examine in detail the theory of credit constraints based on bank lending rules and report the main findings of the empirical literature, pointing out a number of issues raised by the empirical results. In particular, we provide some alternative interpretations for empirical findings that ostensibly seem to suggest the existence of credit constraints. Section 6 deals with search issues associated with the concept of credit constraints. Sections 7 and 8 examine respectively market- and government-led solutions to such constraints in practice. Section 9 summarizes and concludes.

### 1.1. Definitions

*1.1.1. Small Firm* There are many definitions of a small firm, but most rely on the number of employees falling below a certain threshold (Bank of England, 2003). Sometimes this threshold is combined with one for sales. Thus, Berger and Udell (1998), in an important survey of small business finances, define a small firm as one with fewer than 20 employees *and* less than \$2m in annual sales (in 1993 dollars). Other defining characteristics of a small firm include that it is not a subsidiary of another company and hence not controlled by another business, and that it is privately held rather than publicly quoted.

*1.1.2. Debt* Damodaran (1999), in an influential textbook on finance, defines debt in terms of the cash flows associated with it. “A *debt claim* entitles the holder to a contracted set of cash flows (usually interest and principal payments), whereas an *equity claim* entitles the holder to any residual cash flows left over after meeting all other promised claims” (p. 214). This definition is useful because it focuses attention on underlying fundamentals of debt, that is, the fact that it is a *claim*, that it is a claim to the firm’s *cash flows* and that it is *contracted* upon by the firm and bank (or other lender). By contrast, with equity, debt is not a *residual* item payable after other well-defined claims have been met: it has *priority* over certain other claims.

We interpret the term “contracted upon” as debt service payments (usually capital and interest) that can be calculated according to a well-defined rule and based on observable quantities. This allows us to include both overdraft finance and trade credit (borrowing from suppliers or lending to customers) in the discussion.

In order to place bank finance in context, in this chapter we shall discuss the more common forms of debt, namely, trade credit, overdrafts, and term loans (fixed and variable rate). However, our focus will be on bank products, the last two in this list, the two most commonly referred to as debt.

*1.1.3. Credit constraint* A firm is said to be *credit-constrained* in this chapter when the lender (typically a bank) offers less than is socially optimal given the

prevailing rate of interest. The socially optimal amount is that which maximizes one party's (the firm's) objective function subject to the other party (the bank) receiving at least the minimum necessary to remain in business.

### 1.2. *Typical and Sophisticated Small Firms*

Furthermore, an understanding of the role of debt in the small firm and whether small firms can (or indeed wish to) borrow "sufficiently" in financial markets requires a knowledge of the characteristics of the small firm. I will therefore recap a number of well-known facts about the small firm for the reader before launching into the survey proper. In so doing I will make the important distinction between the *typical* small firm and the *sophisticated* small firm since small firms are by no means a homogeneous population.

The *typical* small firm and its entrepreneur(s) statistically has the average characteristics of the population of small firms in a country. It is a micro business (i.e., one with fewer than 10 employees) run by an owner-manager, and its founder is motivated by the desire for independence, to be his/her own boss or to exit from unemployment (Cressy and Storey, 1994). It is not growth-oriented, being stationary in employment terms over long periods (Watson, 1990). Economically, it functions mainly as a source of income for its owners who are mainly self-employed (owners of unincorporated businesses rather than founders of limited companies) and who pay personal tax rates. The owner of the typical small firm has few academic qualifications, though often has work experience in the area of the start-up and, in a significant minority of cases, may have run a business previously. (S)he will often be an asset owner, possessing his own home (Cressy and Storey, 1994). The typical small firm is not very profitable, having many competitors and being rather inefficiently run. Consequently, the firm has a rather brief life, lasting no more than two or three years on average (see below for the U.K. case and Evans and Leighton (1989) for an examination of the U.S. evidence). The typical small firm thus defined accounts for the overwhelming majority of businesses in a country: in Europe and North America it accounts for over 95% of businesses (see ENSR, 1997).

It is clear from the picture we have just painted that the typical entrepreneur has few serious growth ambitions. This makes him (for it is more likely to be a him than a her) unlikely to borrow from the bank (especially at start-up) except for purposes of managing cash flow from daily operations. Thus, about a third of U.K. start-ups in 1988 borrowed for this purpose on Overdraft (U.S.: Line of Credit) rising to one half three years down the line. By contrast, only about 10% of start-ups borrowed to finance the purchase of fixed capital (land, premises, machinery, computers, etc.) (Cressy, 1993). A significant source of borrowing to start a business in fact comes from friends and relatives (U.S.: 'Love money'). Little is known about the structure of this

form of finance but one would expect that the conditions (interest rates and repayment schedules, etc) would be more lenient than those from commercial sources (see Basu and Parker, 2001).

In contrast to the typical small firm we can also identify within the population another stereotype: the *sophisticated* small firm. This animal differs markedly from the average firm.<sup>2</sup> For example, it is estimated that over a ten-year period about 50% of new jobs in the U.K. are generated by only 5% of businesses—these are examples of what we mean by sophisticated businesses (Watson, 1990). Such businesses then are a small minority of firms. An important distinguishing feature of them is that they tend to be limited companies rather than unincorporated businesses (hence, have limited liability and can be bought, sold or bequeathed) and are run by more skilled and qualified entrepreneurs with much more ambitious objectives. These objectives are distinctive particularly with respect to the entrepreneur's target size and growth. Like the typical small business, the sophisticated business will usually serve niche markets and face considerable competition from other firms. However, because they are resourced and managed much more efficiently than the typical small firm, they have higher profit rates and grow significantly faster.<sup>3</sup> This growth invariably necessitates the use of outside finance, particularly bank finance, which is often long term in nature. In a very small minority of cases (1%–5%), the firm will also use formal or informal private equity, defined as funds provided by the sale of shares to outsiders. However, internal finance by way of retained profits is the main source of investment funds and of future growth prospects.

## 2. SOURCES OF SMALL FIRM FINANCE

In this section we discuss the various sources of small business finance together with their primary function and their advantages and disadvantages.

There are two obvious ways of looking at the importance of different sources of finance to the small firm: One is to ask what proportion of the population of small firms use that source at any given stage in their lives; the other is to focus only on the users and to ask, when used, what proportion of total capital the source constitutes. The former criterion refers to the importance *across* firms, the latter to the importance *within* firms. While data for finance usage at all stages in a firm's life seem not to be available, Figure 8-1 from Cressy (1993) shows the importance of various sources of funds for a representative sample of 2000 U.K. firms which started up in 1988. We can see from the chart that on both Across and Within criteria, owner equity (typically personal savings) and bank finance are important (in that order), followed by "love money," that is, funds from friends and relatives. Other sources of finance

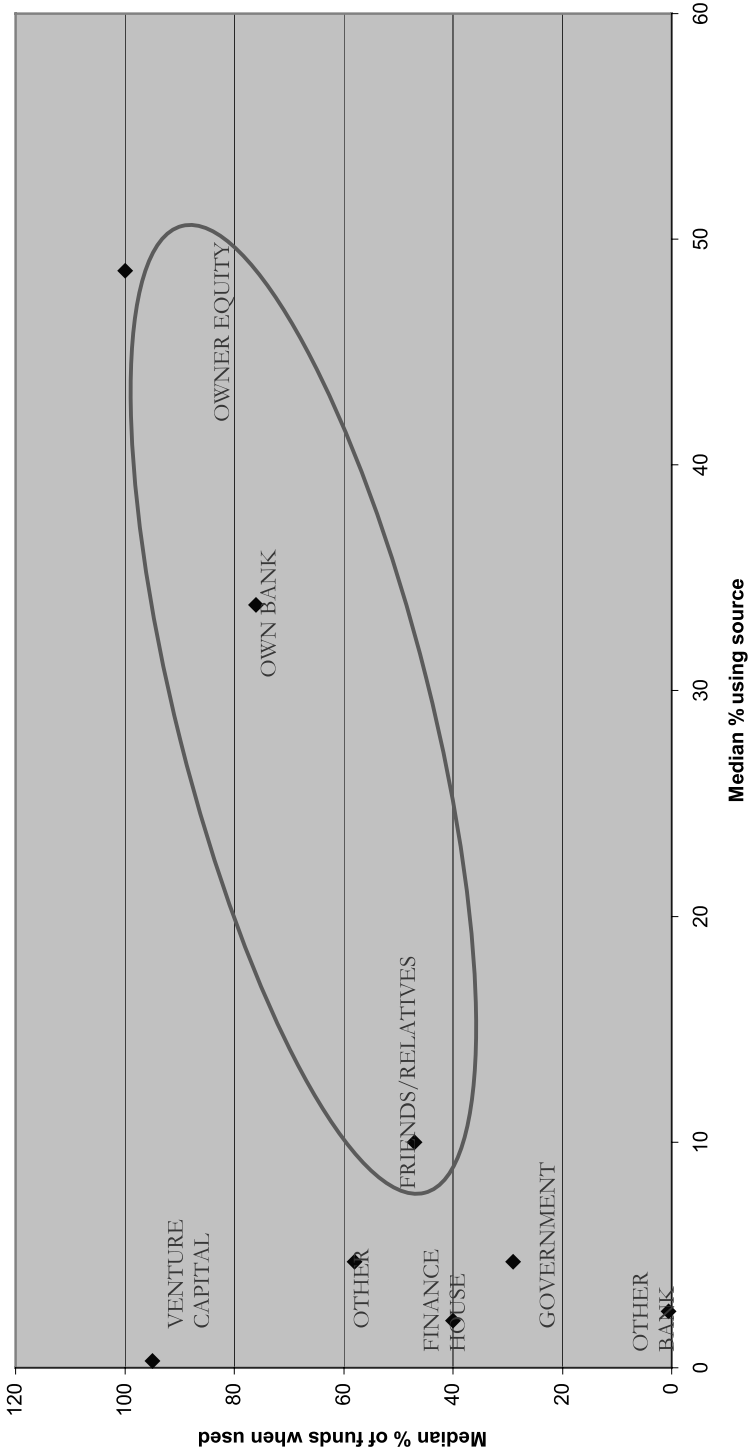


FIGURE 8-1 The frequency and importance of finance sources.



such as venture capital make up a large proportion of funding if used but are not used by (in fact, not available to) the vast majority of businesses.

It is clear then that bank finance is important for the start-up, even though, as already noted, in the U.K. in 1988 only 1/3 of start-up businesses borrowed on overdraft to start their businesses, rising to 1/2 in 3 years. An even smaller percentage of start-ups (about 10%) were financed by a term loan.<sup>4</sup> The picture in the U.S. is similar, as shown in Petersen and Rajan (1994), henceforth PR. PR used a sample of some 3500 small<sup>5</sup> firms from the Survey of Small Business Finances in 1988 and 1989 and found that only about one-third of the smallest firms (defined as those with less than \$15,000 in assets) borrow from any source at all (including banks). This percentage rises steadily with size reaching over 90% for firms with over \$1.3m in assets.

Likewise, they find that overall borrowing by small firms *decreases* with age as borrowers pay back loans from friends and relatives. For example, of very young firms (those less than two years old), 79% borrow, whereas only 59% of older firms (older than 30 years) do. The PR dataset therefore shows that the *overall* borrowing propensity of small firms (defined as the proportion of firms of a given age/size that borrow) declines with age but increases with size. But what of borrowing from banks? While U.S. data appears not to be available, U.K. data suggests that the *bank* borrowing propensity actually *increases* with firm age—at least in the early years (Cressy, 1993).<sup>6</sup> This appears paradoxical but the two propositions are consistent with one another if one allows for the fact that the probability of borrowing from non-bank sources in the States is decreasing at a faster rate than overall borrowing.

The PR dataset also shows that the *fraction* of money borrowed from bank sources (when borrowing is done) rises with size of firm starting at about 50% of borrowed funds for the smallest firms, and rising to over 60% for the largest. The proportion of borrowed funds coming from the bank has, however, an ambiguous relation to age of firm, first rising (from 49% for the youngest firms) and reaching a peak of 63% for firms between 10–19 years before falling back to around 50% for the oldest firms.

In summary, the typical small firm is a reluctant bank borrower and whilst its borrowing propensity will initially increase as the firm ages, bank borrowing eventually becomes less important. However, if the firm grows in size borrowing will become necessary to finance both the fixed capital investment required by expansion and the greater working capital requirements implied by its larger volume of sales.

### 2.1. Short- versus Long-term Liabilities

Small firms are much more dependent on short term than on long term finance. For example, Cosh and Hughes (1994) showed that throughout the U.K.

in the period 1987–89 approximately one third (Services) rising to one half (Manufacturing) of a small firm's liabilities were short term. This contrasts with the constant one third of liabilities of the large firm. This reflects the fact that small firms typically expand very little and that their main source of external funds is trade credit and overdrafts to finance working capital.

## 2.2. Overdraft Finance (U.S.: Line of Credit)

The small firm borrower's first port of call for external funds is, as we have seen, the bank. Furthermore, it is typically a call for an overdraft facility to finance working capital requirements arising from the mismatch of cash inflows from business sales and outflows from business purchases or wage payments rather than a call for a term loan to purchase fixed assets. Being typically located in a service industry, the small firm has less need of fixed assets to expand—at least in the early stages of its development.

So what is an overdraft? An overdraft is a facility offered by the bank that allows the firm to borrow up to a maximum (the Overdraft *Limit* in the U.K.) during a fixed time period at a margin determined in advance. The interest rate paid by the firm is therefore the sum of the margin and Base or Prime rate and therefore varies with the latter. Hence, it is expressed as Base rate plus  $x$ , where  $x$  is the margin charged.<sup>7</sup> Payments of interest on an overdraft in the U.K. are calculated on a daily basis on the drawn-down amount at the beginning of that day and must be paid within the period of trade credit, for example, three months. However, the actual payment schedule is flexible—within informal limits. This indeed is one of the attractive features of the overdraft to small businesses whose cash flows are often highly variable and the reason why it will typically be used in preference to a term loan (see below). An overdraft has another important feature not widely known, namely that it is in principle repayable *on demand*. However, the bank is unlikely to enforce this legal right except in situations where the repayment of the overdrawn amount is in jeopardy. Figure 8-2 illustrates a potential rationale for the use of an overdraft.

For simplicity assume that the firm has purchases and sales starting at time 0 and that these grow linearly from a pair of initial values so that net sales  $S(t)$  also grow linearly. Sales and purchases both last for a period of  $t^* =$  three months and then cease but cash is now starting to come in steadily. In six months time ( $t = 2t^*$ ) the firm (for simplicity) closes. The firm can delay payments for purchases from suppliers for three months but by the same token will receive payments from customers only three months later.<sup>8</sup> However, to produce output a workforce is required. Wage payments of  $w = 1$  per week therefore mean that the firm has a negative weekly gross cash flow ( $-w = -1$ ) in the absence of a bank. This need to pay wages from the start means that,

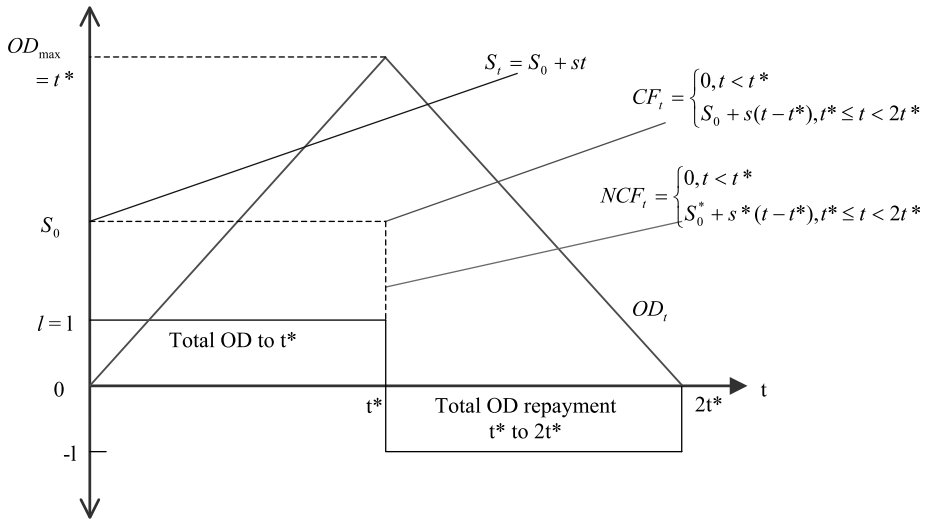


FIGURE 8-2 A motive for the overdraft: timing of the firm's cash flows.

provided the bank will offer the facility, the firm is steadily accumulating an overdraft at a rate of  $l = 1 = w$  per day until it reaches  $t^* =$  three months, yielding maximum drawdown of  $OD_{max} = lt^* = t^*$  in Figure 8-2. Then cash starts coming in steadily and the overdraft can be paid back (for simplicity) at rate  $w$  per day eventually reaching zero at  $2t^* =$  six months.<sup>9</sup> Its cash flow  $CF(t)$  and net cash flow  $NCF(t)$  ( $CF$  net of overdraft repayments) are also shown in Figure 8-2. Note that the firm's net present value can be positive over the six month period, but without the overdraft facility (or something equivalent) it would not even start in business.

How important are overdrafts? Cosh and Hughes (1994) report that in the U.K. overdrafts finance on average are about 11% of a small firm's assets (about the same for Manufacturing and Services) which is approximately twice the proportion financed by overdrafts in large manufacturing firms (at 6%) and almost three times the proportion financed by overdrafts in large service industry firms (4.4%).

What explains the form of overdraft contracts? Cressy (1995) used a simultaneous equations approach to the analysis of overdraft lending to U.K. start-ups. I modeled the determinants of overdraft limits in which the limit, interest margin, security and survival were all endogenous to the system. Exogenous variables included age of the entrepreneur, team size, experience and industry. Estimating this on a random sample of 2000 U.K. start-ups in 1988, I found that (a) margins declined with the overdraft limit (probably for reasons of economies of scale in administration); (b) security required for an overdraft increased with the size of the limit; (c) security placed on a loan increased with

the age of the entrepreneur (perhaps reflecting both availability of collateral—higher for older people as their wealth is higher<sup>10</sup>—and the greater chances of survival—older entrepreneurs last longer in business). Thus human capital and assets (both correlated with business survival and hence with the continuation of the bank account) appear to be important determinants of whether and how much money the entrepreneur received from the bank.

### 2.3. Trade Credit

Trade credit is a form of borrowing whereby the firm will purchase raw materials (or, in the case of retailers, finished products) but pay for them later. Thus, if Minnow Ltd buys £10,000 of goods from Gudgeon Ltd and receives them immediately but pays for them in three months time, it has effectively received a *deposit* from Gudgeon of £10,000 for three months. The interest rate applicable is the opportunity cost of funds. In other words, if the interest rate on an overdraft is 10% per annum, then Minnow Ltd is effectively *receiving* a quarterly rate of interest of 2.4% and Gudgeon Ltd is effectively *paying* the same rate on the £10,000. Trade credit (net debtors) is by far the most important source of funding for a small firm's assets in the U.K. and is more important for small than large firms. Cosh and Hughes (1994), for example, show that approximately a third of a small manufacturing firm's assets are funded in this way compared to a quarter of a large firm's. In service industries, their data show that the large firm funds even less from trade credit: approximately one fifth of its assets are paid for this way.

Trade credit in the U.S. is somewhat more important to the small firm than in the U.K., constituting 16% of total debt in 1992 (see Berger and Udell, 1998, and Table 8-1). While figures are not available for the U.K. or other parts of Europe, in the U.S. the *proportion* of total finance accounted for by trade credit varies little with the age of the firm, remaining roughly constant at 13–14% throughout most of the age range (though rising “briefly” to a peak of 17% in the 5–24 age group). It is also significantly higher for large firms (17%) than small (12%) possibly reflecting the large firm's greater market power over suppliers, which are often themselves smaller firms.

The U.S. data also enable us to see how bank debt in the form of a line of credit varies with trade credit, age and size. Table 8-1 shows that the absolute amounts of trade credit and bank debt rise until firm “maturity” at about 25 years and tail off in “old age,” reflecting growth to maturity and possibly stagnation thereafter. Table 8-1 shows moreover that trade credit is no less than *five times* as important for large as small firms (reflecting their higher turnover) and for bank debt the corresponding ratio is about *four* times (reflecting the need to finance cash flows associated with the larger volume of trade credit). Thus, the interrelationship between trade credit and bank lines of credit seems

TABLE 8-1 Sources of small firm finance

|   | Sources of equity |               |                 |              | Total equity |
|---|-------------------|---------------|-----------------|--------------|--------------|
|   | Principal owner   | Angel finance | Venture capital | Other equity |              |
| A: All nonfarm, nonfinancial, nonreal-estate small businesses | 31.33%            | 3.59%         | 1.85%           | 12.86%       | 49.63%       |
|   | \$524.3           | \$60.0        | \$31.0          | \$215.2      | \$830.6      |
| B: Breakout by size of small business                         |                   |               |                 |              |              |
| “Smaller”   | 44.53%            | NA            | NA              | NA           | 56.00%       |
| (<20 empls<br>& <\$1m sales)                                  | \$175.7           |               |                 |              | \$220.9      |
| “Larger”  | 27.22%            | NA            | NA              | NA           | 47.67%       |
| (≥20 empls<br>or ≥\$1m sales)                                 | \$348.1           |               |                 |              | \$609.6      |
| C: Breakout by age of small business                          |                   |               |                 |              |              |
| “Infant”  | 19.61%            | NA            | NA              | NA           | 47.9%        |
| (0–2 yrs)   | \$8.6             |               |                 |              | \$21.1       |
| “Adolescent”  | 17.37%            |               |                 |              | 39.37%       |
| (3–4 yrs)   | \$25.1            |               |                 |              | \$56.8       |
| “Middle aged”   | 31.94%            | NA            | NA              | NA           | 48.00%       |
| (5–24 yrs)  | \$324.9           |               |                 |              | \$488.2      |
| “Old”   | 35.42%            | NA            | NA              | NA           | 56.5%        |
| (25 or more yrs)  | \$165.8           |               |                 |              | \$264.5      |

to be potentially quite strong, though I know of no study that examines this relationship systematically controlling for other relevant factors.

#### 2.4. Term Loans

A term loan is a loan for a fixed amount with a fixed duration and a regular payment schedule decided and contracted upon at the outset. The rate charged may be fixed or variable. A fixed rate implies the interest rate is independent of the base or prime rate prevailing. A variable rate by contrast implies that the margin above base is fixed but the interest rate (sum of base and margin) varies with the base rate. The purpose of a term loan can vary widely but it is usually for the purchase of land or buildings, machinery, improvements to buildings and so on. In other words, it is usually used to finance fixed- or long-term capital rather than working or short-term capital.

For a *repayment* loan, where the principal as well as the interest on the loan is part of the scheduled payment, the formula for the periodic payments is given by an inversion of the familiar *annuity* formula (see Brealey and

TABLE 8-1 *Continued*

| Sources of debt   |                   | Nonfinancial business and government |              |                  |              |                 |             |                  |            |                        |  | Total equity plus debt |  |
|---|-------------------|--------------------------------------|--------------|------------------|--------------|-----------------|-------------|------------------|------------|------------------------|--|------------------------|--|
|   |                   | Financial institutions               |              |                  | Trade credit |                 |             | Other businesses |            | Govt.                  |  | Individuals            |  |
| Commercial banks  | Finance companies | Other fin. insts                     | Trade credit | Other businesses | Govt.        | Principal owner | Credit card | Other indivs     | Total debt | Total equity plus debt |  |                        |  |
| A: All nonfarm, nonfinancial, nonreal-estate small businesses |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 18.75%  | 4.91%             | 3.00%                                | 15.78%       | 1.74%            | 0.49%        | 4.1%            | 0.14%       | 1.47%            | 50.37%     | 100.00%                |  |                        |  |
| \$313.8   | \$82.1            | \$50.1                               | \$264.1      | \$29.2           | \$8.1        | \$68.5          | \$2.4       | \$24.5           | \$842.9    | \$1,673.4              |  |                        |  |
| B: Breakout by size of small business                         |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| "Smaller"   |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 14.88%  | 3.08%             | 3.53%                                | 11.81%       | 1.06%            | 0.37%        | 5.59%           | 0.53%       | 3.16%            | 44.00%     | 100.00%                |  |                        |  |
| (<20 empl   | \$58.7            | \$12.1                               | \$13.9       | \$46.6           | \$1.4        | \$22.1          | \$2.1       | \$12.5           | \$173.6    | \$394.5                |  |                        |  |
| & <\$1m sales)  |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| "Larger"  |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 19.94%  | 5.47%             | 2.83%                                | 17.01%       | 1.95%            | 0.52%        | 3.63%           | 0.02%       | 0.94%            | 52.33%     | 100.00%                |  |                        |  |
| (≥20 empl   | \$255.0           | \$70.0                               | \$217.5      | \$25.0           | \$6.7        | \$46.5          | \$0.3       | \$12.0           | \$669.3    | \$1,278.9              |  |                        |  |
| or ≥\$1m sales)   |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| C: Breakout by age of small business                          |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| "Infant"  |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 15.66%  | 8.33%             | 3.84%                                | 13.4%        | 1.52%            | 0.33%        | 6.04%           | 0.21%       | 2.77%            | 52.1%      | 100.00%                |  |                        |  |
| (0-2 yrs)   | \$6.9             | \$3.7                                | \$5.9        | \$0.7            | \$0.1        | \$2.7           | \$0.1       | \$1.2            | \$23.0     | \$44.1                 |  |                        |  |
| "Adolescent"  |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 30.84%  | 2.51%             | 2.36%                                | 13.42%       | 1.06%            | 0.72%        | 6.19%           | 0.2%        | 3.32%            | 60.63%     | 100.00%                |  |                        |  |
| (3-4 yrs)   | \$44.5            | \$3.6                                | \$19.4       | \$1.05           | \$1.0        | \$8.9           | \$0.3       | \$4.8            | \$87.4     | \$144.2                |  |                        |  |
| "Middle aged"   |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 17.86%  | 5.85%             | 2.87%                                | 17.10%       | 2.39%            | 0.44%        | 3.91%           | 0.17%       | 1.42%            | 52.00%     | 100.00%                |  |                        |  |
| (5-24 yrs)  | \$181.6           | \$59.5                               | \$174.0      | \$24.3           | \$4.4        | \$39.7          | \$1.7       | \$14.4           | \$528.9    | \$1,017.1              |  |                        |  |
| "Old" (25 or more yrs)  |                   |                                      |              |                  |              |                 |             |                  |            |                        |  |                        |  |
| 17.25%  | 3.28%             | 3.38%                                | 13.86%       | 0.56%            | 0.54%        | 3.68%           | 0.06%       | 0.88%            | 43.5%      | 100.00%                |  |                        |  |
| \$80.8  | \$15.3            | \$15.8                               | \$64.9       | \$2.6            | \$2.5        | \$17.2          | \$0.3       | \$4.1            | \$203.6    | \$468.0                |  |                        |  |

Estimated distributions of equity and debt, percent or total equity plus debt (top numbers) and billions of dollars (bottom numbers). Source: Berger and Udell (1998).

Myers, 2003):

$$L = \frac{C}{1+r} + \frac{C}{(1+r)^2} + \cdots + \frac{C}{(1+r)^T}$$

$$= \frac{C}{r} \left[ 1 - \frac{1}{(1+r)^T} \right],$$

where  $L$  is the loan amount,  $C$  is the constant per period payment of capital and interest,  $T$  is the duration of the loan and  $r$  the rate of interest charged on the loan. Solving this for  $C$  gives:

$$C = \left[ 1 - \frac{1}{(1+r)^T} \right]^{-1} Lr.$$

(For example, with a one-period loan we find that  $C = L(1+r)$  so that the total amount of the loan is repaid in one go, and the interest only accumulates for one period; and for an infinitely long-lived loan we find that  $C = Lr$ , and only interest, and not capital, is paid, but indefinitely.) It is not difficult to show that the payments will be increasing in the loan amount and in the interest rate, and decreasing in the duration of the loan, other things held constant.

Small firms have a smaller proportion of loans from long-term sources, that is, typically term loans.<sup>11</sup> In Cosh and Hughes' (1994) data we find that only about 3% of total capital and liabilities of small manufacturing firms in the U.K. were funded by long-term loans whereas 11.5% of large firms' capital and liabilities were so funded. Almost the same ratio of 4:1 for small and large firms applies to service-based companies. As mentioned above, this mirrors the rate of replacement and funding of growth opportunities of the two sizes of firm.

### 3. SECURITY FOR LOANS: COLLATERAL

Collateral or security for a loan is an important feature of small business lending. Naturally, not just any asset is appropriate as collateral for a bank. Banks are suspicious of intangibles, things that can't be touched or seen and of items with short or passing lives.<sup>12</sup> So banks are likely to ask as the basis for security tangible (rather than intangible) and fixed (rather than current) assets. Classic examples of tangible fixed assets would be land, plant and buildings. Contrast this with intangible assets such as patents and copyrights. A computer program would typically be protected by a copyright (since it is lines of *written* code in which the author has a *right to prevent unauthorized copying*). The code might be worth literally millions of dollars but its value is often difficult to certify and therefore would not typically form the basis of collateral. Likewise,

the patent on a new drug would not typically be suitable for this purpose. This again might be highly valuable, but the bank would tend to regard it as problematic from the point of view of salability.

Land, machinery and buildings are tangible, fixed assets. They have relatively long lives and larger used values, particularly in the case of land and houses located in “good” residential areas. Current assets by contrast would be items such as stocks of goods and work in progress, debtors and current interest earned. These might in the event of default have zero value. Debtors might be valueless if the company to whom the products or services were sold, and who currently owed money for them in return, went bust. Half-finished goods are of little worth on the open market and even finished goods may need to be sold in distressed conditions. Hence, typically little weight is placed on them in deciding the value of collateral.

### *3.1. Economic Rationale for Collateral*

There are several reasons for the placing of collateral identified in the theoretical literature (see Leeth and Scott, 1990; Coco, 2000). However, relatively few of these hypotheses have been tested empirically. Motives for the use of collateral include:

- Reducing agency costs—e.g., by preventing asset substitution when the borrower sells the collateral pledged and invests it in a high risk project.
- Reducing debt expense—by allowing the lender title to specific assets in the event of default so avoiding costly fights among creditors over who gets which assets.
- Reducing the cost of debt—by mitigating the problem of claim dilution which occurs when subsequent debt is issued with higher priority.

This theory produces a number of testable implications. In particular, collateral will be more likely to be placed:

1. The lower the percentage of specialized assets<sup>13</sup>—because firms with such assets will have less incentive to substitute high for low risk projects.
2. The greater the duration of the loan—e.g., long term ‘reputation’ losses are higher with a short term loan compared with the immediate gains from ‘ripping off’ the bank.
3. The greater the size of loan—e.g., the fixed bank administrative costs of “perfecting” collateral per Euro lent decline with loan size, making larger loans more likely to have a collateral requirement.



4. The higher the risk free interest rate—e.g., higher real interest rates raise debt payments by more than operating cash flows and decrease the firm's interest coverage; hence an increase in the real rate of interest expands the use of secured debt.

The theoretical relationship of collateral provision to borrower quality, arguably the most interesting issue in the theory of collateral, is unfortunately ambiguous. On the one hand, the theory of collateral as a means of *sorting borrowers by observed risk* (Berger and Udell, 1990) argues that banks reduce their risk exposure by requiring collateral from more risky borrowers. This predicts a *negative* relationship between borrower quality and collateral with lower quality borrowers placing more. On the other hand the theory of *collateral as a screening device* argues that collateral enables low risk borrowers to be separated from high risk borrowers by the high risk borrower placing more collateral (Boot, Thakor and Udell, 1991). This generates a *positive* relationship between collateral and firm risk.

So what of the empirical evidence? Leeth and Scott (1990) in an early study of collateral determinants used samples of 12,000 and 14,000 established small businesses in the U.S. for the years 1980 and 1982 to test the five hypotheses above. They estimated a probit model (where the dependent variable took values of one if the loan is collateralized; and zero if not) using proxies for the variables identified above and found that the predictions of the model were by and large borne out in the data. In particular, they found (as did a later study by Berger and Udell) that *more risky loans were more likely to be collateralized* supporting the sorting by observed risk hypothesis. The only variable with no significant effect, but which should have carried a positive sign, was the proxy for the safe rate of interest. (However, they argue that its statistical non-significance may have been due to problems in defining the proxy itself rather than representing a falsification of the theory). Recent U.K. research using a more general model of lending contracts has found that the evidence supports the sorting by screening rather than sorting by observed characteristics hypothesis (Cressy and Toivanen, 2000; Han, 2005).

Recent U.K. research has probed both the relationship between collateral and borrower quality and the information regime in which it is embedded (Cressy and Toivanen, 2001, and Toivanen and Cressy, 2000). These studies test the role of collateral as a signaling mechanism (a regime of adverse selection—see below) versus its role as an incentive mechanism (a regime of moral hazard—ditto). But both motives exist only in an environment of asymmetric information, specifically where the firm knows more about the quality of its project or the amount of borrower effort than the bank. Thus it is essential to test for the nature of the information regime, and conditional on asymmetric information being revealed, for the form of asymmetric information.

In the case of adverse selection, an optimal contract will distinguish borrower types (given) and have a “good” (high quality) borrower placing more collateral (and getting a lower interest rate in return). This separates her from the “bad” borrower who will try to dissemble her type and get a better contract (lower interest rate). This screening function of collateral yields, as we have mentioned, a *positive* relation between borrower quality and collateral provision, with collateral the “independent” variable. By contrast, in the case of moral hazard, project quality (probability of success or of loan repayment) is *endogenous* and causality runs from the placing of collateral to the success of the project: more collateral induces more effort (since collateral is exchanged only if the project fails) yielding the *positive* relationship between collateral provision and success (effort) with collateral the “independent” variable.

### 3.2. Lending, Collateral and Personal Guarantees

Using the well-established technique for reducing risks by matching assets and liabilities, banks may also be prepared to offer overdraft facilities using as collateral current assets like debtors (accounts receivable). This is because the duration of the overdraft matches the duration of the collateral and the receipts from these sales are legally valid documents which, as a last resort, enable payment to be extracted from the purchaser by court action. On the bank’s asset side, overdrafts are, in principle, repayable on demand, in other words, within 24 hours notice from the bank. Therefore, the risk associated with debtors used as collateral for lines of credit is not great. This is particularly true if (as is often the case) the sales in question are to a large firm.

Personal guarantees are a common feature of small business lending and frequently negate the principle of limited liability which states that an owner of a limited company is liable for the company’s debts only up to the value of his/her shares. With personal guarantees, the owners are made personally liable for the company’s debts and their limited liability evaporates. Despite the frequency of such guarantees (Han, 2005, reported that in the U.S. in 1998, 55% of small business loans were subject to them and almost one third of loans required both guarantees and collateral) most of the theoretical literature on lending to business firms (small or large) is predicated on the assumption of *limited* liability.

Because of the considerations discussed thus far, banks will often not lend to very small, young businesses without collateral in the form of fixed assets or personal guarantees based on personal assets (e.g., the owner’s house). For more established customers where risk can be assessed, banks will lend without collateral and raise the interest rate to compensate for the extra risk. This may, in theory, itself create something of an adverse selection problem,

as higher interest rates may attract lower quality borrowers, see Stiglitz and Weiss (1981), potentially leading to credit rationing by the bank. However, there is very little evidence to show that this is a problem in practice.

#### 4. TERM LOAN FEATURES

What are the main features of actual long term bank loan contracts to small firms? There is only limited information on this subject at present.

Cressy and Toivanen (2000) provided data on the contract terms to some 2800 term loans to U.K. small businesses (both new and established) made during the period 1987–1993. The average loan size was a modest £19,000. Collateral or loan security was provided in the majority (62%) of cases. Margins were on average 3% above Base or Prime rate and the loan had an average duration of 7–8 years. Interestingly, only about 8% of the loans defaulted during the term of the loan despite the fact that the period of the study encompassed a severe recession.<sup>14</sup>

Leeth and Scott (1989) (LS), using survey data from the National Federation of Independent Businesses, examined the frequency and determinants of collateral provision including the type of collateral required/placed and underlying assets used. LS also report that 60% of loans are secured by some form of collateral, whether business or personal. However, the type of collateral varied. For example, 30% of loans offered business collateral only, 10% personal collateral only and 20% both types of collateral. Regarding the underlying assets, plant and equipment was the most frequent asset for business collateral and housing assets the most frequent for personal.

Petersen and Rajan (1994) using a large sample of US small businesses identified risk factors that influence the interest rate paid on a loan. Borrowers in lower default premium classes together with larger and older borrowers (the latter measured by assets) received lower interest rates. This reflects their lower credit risk.<sup>15</sup> Likewise, borrowers with greater interest cover (defined here as the ratio of profits to interest payments) received lower margins over prime, reflecting their greater debt servicing capacity and lower bankruptcy risk.

Finally, Cressy and Toivanen (2000), in a structural simultaneous equations model (see below) found, on the one hand, that collateral provision lowered the interest rate charged on U.K. term loans and, on the other hand, that larger and longer duration loans were more likely to be collateralized. Collateral was also positively correlated with success, consistently with the moral hazard-reducing role of collateral identified above.

#### 4.1. Explanations for Term Loan Features

Earlier we examined earlier some of the empirical reasons for collateral provision in lending contracts. However, other parameters that deserve discussion include loan availability, loan size and interest margins.

Early work by Bates (1990) on a sample of some 4500 firms owned by nonminority U.S. males suggested a correlation between personal characteristics and credit availability. In Bates' words, "a highly educated white middle aged male who is investing large sums of equity capital in his small business is going to have maximum access to debt capital." Later Cressy (1996), using a sample of 2000 U.K. start-ups in 1988, found that the characteristics that predicted small business survival had much in common with the characteristics that determined bank lending. Thus owner age, team size, work experience in the area of the start and whether the business was a purchase predicted both survival and loan size. However, both of these studies used single equation methods (logit) and neither dealt satisfactorily with the potential endogeneity problems associated with the data.

Theorists have spent considerable time analyzing the contractual aspects of lending under asymmetric information. Some features of interest arising from that analysis are relationships between collateral, loan size, interest rates and default. Such theories suggest that contract features reflect the bank's attempt to deal with the problems of adverse selection and moral hazard (see the survey by Parker, 2002). Agency problems (e.g., lazy borrowers) and adverse selection problems (unknown borrower quality) between the bank and the firm can, in theory, be mitigated by the use of collateral either as a motivating device (it is lost only under default) or as a separating device (better quality borrowers will place it).

The empirical evidence on this front has lagged far behind the theoretical advances, although some empirical studies exist. Toivanen and Cressy (2000), for example, provided both an encompassing model of small business term lending and an empirical test on a sample of some 2800 U.K. small business loans. They examine not merely the role of asymmetric information but also of the bargaining power of the firm and bank in determining loan contract structure. Their main findings are that:

- (a) The information regime is asymmetric rather than symmetric, thus opening the gates to either moral hazard or adverse selection.
- (b) The bank is, in practice, more concerned about moral hazard (borrower "laziness") than adverse selection (attracting the "wrong" kind of borrowers) and designs contracts to deal with this issue in mind. This is shown by the positive relationship between the probability of default and the collateral level of the contract and the fact that default rates are endogenous (i.e., determined by unobservable borrower effort levels).

- (c) The bank deals with moral hazard by imposing collateral requirements on loans (to enhance unobservable borrower effort<sup>16</sup>).
- (d) More collateralized loans are less risky and so get lower margins.
- (e) Larger loans are offered to better borrowers (those with lower default rates).

## 5. CREDIT CONSTRAINTS

So far, this chapter has examined the usage and characteristics of debt in the small firm balance sheet. An important question, seemingly highlighted by the “opaqueness” of the small firm to outside financiers, relates to the *availability* of finance to the small and young firm: Are small firms credit-constrained by the financial system?

It would be nice to be able to give a straight answer to this question, and one based firmly on empirical evidence. Unfortunately, this is currently still not possible.<sup>17</sup> Despite this fact, the debate over the issues is fascinating.

Let us begin by defining carefully what we mean by a credit constraint. We say that a credit constraint exists if at least on firm is unable to obtain the correct amount of finance for a viable project. “Viable” means one with a *positive net present value*. Often credit constraints will assumed to be related to bank lending policy which relies on collateral, and therefore the assets of the borrower. However, since assets must to some degree depend on the human capital of the entrepreneur, human as well as financial capital will be likely to be implicated.

The theoretical literature regarding credit constraints on small businesses is, as we have mentioned, huge (see Parker, 2002, for an excellent survey). The empirical literature is, by contrast, much smaller but growing rapidly. Particular areas of academic and policy interest center on the following questions:

- (a) How realistic is it to assume that wealth or human capital of entrepreneurs is exogenous (outside his/her control)?
- (b) Are the criteria for credit constraints defined properly?
- (c) Are there other plausible explanations than credit constraints for what we identify empirically as credit constraints? (E.g., entrepreneurial risk aversion or control aversion)?

We discuss each of these questions in turn.

### 5.1. Switching and Credit Constraints

One of the biggest areas of empirical research focuses on the factors determining whether an individual switches into or out of self-employment (SE). This is particularly important in view of the role of SE in generating jobs in the modern economy. It is also considered to be an area where credit constraint are most likely to explain any apparent deficiency (excess) of numbers entering (leaving) SE.

Evans and Jovanovic (1989) (henceforth EJ), in a now celebrated paper, developed a theory of credit constraints based on the idea that banks lend in proportion to a firm's assets rather than on the basis of its expected cash flow profits.<sup>18</sup> The result may be that there is insufficient lending and excessive failure of cash-starved businesses. The formal model is worth presenting since it or its variants have been subject to a number of empirical tests. It runs as follows.

An individual can be either a wage earner or an entrepreneur. If a wage earner his/her income is given by

$$w + rz, \quad (1)$$

where

$$w = \mu x_1^{\gamma_1} x_2^{\gamma_2} \quad (2)$$

is the expected wage rate conditional on previous experience as a wage worker ( $x_1$ ) and education  $x_2$ .  $\mu$  is a positive constant.  $z$  is the individual's assets invested at interest rate  $r - 1$ . Expected gross income from entrepreneurship is given by

$$y = \theta k^\alpha, \quad (3)$$

where  $\theta$  is the individual's ability as an entrepreneur,  $k$  is her start-up capital and  $\alpha$  is a positive fraction. We note in passing a counterfactual here: human capital denoted by  $x_i$  in (2) is *not* transferable into self-employment and, likewise, entrepreneurial ability  $\theta$  does *not* influence the wage. We shall later provide evidence that this assumption is false.

Since the entrepreneur's assets net of start-up capital are available in business as well as in wage employment he/she continues to earn income from her remaining assets after start-up. The total net income from self-employment is therefore:

$$y + r(z - k). \quad (4)$$

In the EJ model the individual is subject to a banking constraint, namely that a person can borrow only in proportion to one's assets, the borrowing factor being  $\lambda - 1$ ,  $\lambda > 1$ . The amount of capital that can be used in a business is thus constrained by how much one has now ( $z$ ) and how much one can borrow ( $(\lambda - 1)z$ ). The total start-up capital cannot therefore exceed:

$$(\lambda - 1)z + z = \lambda z. \quad (5)$$

The individual first chooses  $k$  to maximize income from SE:

$$\max_{k \in [0, \lambda z]} [\theta k^\alpha + r(z - k)] \quad (6)$$

and then chooses the maximum of the returns to being employed and being self-employed. This yields a demand for capital of

$$k = \left( \frac{\theta \alpha}{r} \right)^{1/(1-\alpha)}. \quad (7)$$

However, if one's demand for capital exceeds  $\lambda z$ , he/she will be credit constrained. Thus unconstrained entrepreneur's ability must be less than  $\theta^*$  defined by

$$\theta^* = (\lambda z)^{1-\alpha} r / \alpha. \quad (8)$$

Entrepreneurs with  $\theta$  less than  $\theta^*$  have an expected income that does not depend on the entrepreneur's assets,  $z$ ; those that breach (8) have an income which increases in  $z$ , since they can only borrow more if their asset or borrowing constraint is relaxed.

The situation is represented graphically in Figure 8-3. We can see that individuals with low  $\theta$  will tend to be wage workers since their income from SE is relatively low. Those with higher  $\theta$  are more likely to be entrepreneurs, but if their ability is "too high" they will be more likely to find credit constraints limit their activity.

The effects of credit constraints should be obvious by now: apart from reducing the scale of the business to a suboptimal level, by the same token they will reduce its profitability, and for the marginal entrepreneur, induce him or her to switch back into wage employment.

EJ argued that empirically we should find credit constraints to SE *if and only if* there is an empirical correlation of assets and switching into SE (or equivalently between assets and SE survival). This is based on the idea of a bank lending rule discussed above. A relaxation of the lending rule (or equivalently an unanticipated increase in fixed assets) will, if businesses are credit constrained,

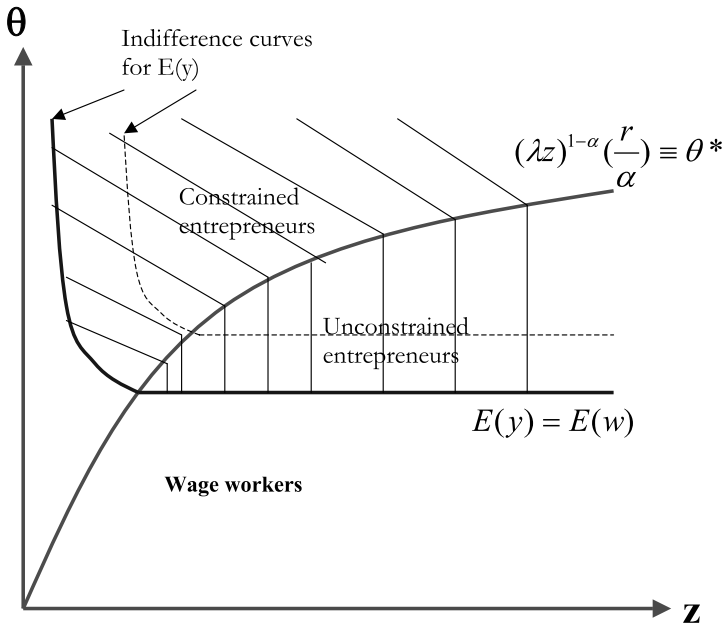


FIGURE 8-3 *Assets, entrepreneurial ability and credit constraints à la Evans and Jovanovic.*

increase switching into SE and increase business survival rates. EJ estimated their model on a sample of 1949 American white males aged between 14 and 24 years in 1966 who were wage workers in 1976 and who were either wage workers or self-employed in 1978.<sup>19</sup> These individuals were between the ages of 24 and 34 in 1976, the typical age of entrepreneurial entry. About 4% of those who were wage earners in 1976 switched into SE by 1978.

EJ estimated the probability of SE as a function of assets (and its square), wage experience, education starting wage income and controls. The coefficient on assets was found to be positive and significant (at the 2% level). On the assumption of zero correlation of assets and entrepreneurial ability, they concluded that liquidity constraints exist. They also estimated SE earnings as a function of the same variables and found that wealthier individuals earn more in SE because they “will have started businesses with more efficient capital levels” (p. 820). Finally, they found that people with smaller assets would be forced to devote a larger proportion of their wealth to their businesses.

### 5.2. Questioning the EJ Result

While a growing number of studies have apparently supported the EJ finding of credit rationing (see, e.g., Holtz-Eakin, 1994a, 1994b; Blanchflower and Oswald, 1998) there are questions about a number of features of their study



that have in turn led to extensions of the model and more sophisticated tests of its hypotheses:

- (a) How appropriate is the model EJ used? In particular are assets endogenous to the system or, for example, a function of the human capital of the entrepreneur, thus making the latter the primary constraint (Cressy, 1996; Astebro and Bernhardt, 2003; Parker and Van Praag, 2004)?
- (b) How should one interpret the EJ results, namely, the positive correlation of assets and survival? For example, it has been questioned whether the repeated findings of various studies can be explained by the existence of uncontrolled factors such as control aversion or risk aversion of the would-be/actual entrepreneur (Cressy, 1995, 1998).
- (c) How representative are the data sets used in the replication studies? For example, in the case of the U.S. studies by Holtz-Eakin et al. (1994b), these authors examined only the top end of the wealth distribution and found evidence supporting EJ. Recent evidence suggests however that the pattern may be radically different through most of the wealth spectrum (Hurst and Lusardi, 2004).

*5.2.1. The EJ Theory Generalized* To clarify the theory behind this (and other issues), we now sketch out a general model of the switching decision in which human capital can play a role in the returns to both wage employment and to entrepreneurship and where it can therefore influence wealth levels of would-be entrepreneurs. This allows us to test whether human capital relaxes potential financial capital constraints by generating collateralizable assets or increases them by increasing the demand for capital at a fixed level of assets.<sup>20</sup>

We can write the return functions in employment and self-employment, respectively, as follows:

$$w = w(x), \quad w' > 0 \quad (9)$$

and

$$y = y(\theta, x, k), \quad y_\theta, y_x > 0, \quad (10)$$

where  $x$  is human capital,  $k$  is financial capital and  $\theta$  is once again entrepreneurial ability. Note that the returns to wage employment and self-employment now both depend on human capital  $x$ . This implies an optimal demand for entrepreneurial capital  $k$  of the form

$$k = k(z, x, \theta), \quad k_z \geq 0, \quad k_x, k_\theta > 0, \quad (11)$$

where wealth,  $z$ , enters the function *only if* the entrepreneur is credit constrained. We may also specify  $z$  as a function of  $x$  and  $\theta$ :

$$z = z(x, \theta), \quad z_x, z_\theta > 0. \quad (12)$$

If older entrepreneurs are more capable entrepreneurs (i.e., have higher  $\theta$ ), the EJ model says that such individuals are more likely to be credit constrained ( $k > \lambda z$ ) because their demand for credit (from (7)) is larger. However, if older borrowers also have more human capital, the result does not necessarily hold since from (12) both  $k$  and  $z$  now increase with  $x$ .

*5.2.2. Endogenous Constraints?* A deficiency of the EJ model, then, is that it fails to allow that human capital may be productive in *both* wage and self-employment and that it may influence the level of assets of the aspiring entrepreneur. Cressy (1996) argued on the basis of U.K. start-up data that once human capital is included in the survival equation the role of assets disappears. In other words, the correlation between assets and survival is simply spurious and due to the simultaneous correlation of human capital (measured by average age of entrepreneur, team size, and work experience in the area of the start-up) with both these variables. This is borne out in his sample of U.K. start-ups. However, more recent studies on different (mainly U.S.) datasets have found that the relationship between assets and switching/survival is weakened but not eliminated once human capital is included.

Astebro and Bernhardt (2003) (AB henceforth) started by generalizing the EJ model incorporating the innovation of Xu (1998) (which allows individual savings to influence wealth) and allowing human capital to influence capital demand and assets. There are two periods as in Xu (1998). All agents start in period 1 with zero wealth but receive a wage income  $w$ , which funds consumption and savings:

$$z = c^* w,$$

where  $c^* = 1 - c^{**}$  is the optimal (exogenous) consumption of the agent as a proportion of her income in period 1. The wage equation is now written as a function of human capital  $x$ :

$$w = \mu x^\gamma,$$

where  $\gamma$  is a positive fraction and  $\mu > 0$ . Entrepreneurial choice is made in period 2. The period 2 income for a wage worker is  $w + rz$  where  $r$  is the rate of interest and  $z$  the worker's wealth. The period 2 profit for an entrepreneur is

$$\pi = y + r(z - k),$$

where  $y$  is self-employment earnings and  $k$  is start-up capital. Earnings in entrepreneurship now also depend on human capital:

$$y = \theta x^\delta k^\alpha,$$

where  $\delta$  and  $\alpha$  are both positive fractions. As in EJ it is assumed that the individual can borrow from the bank up to a proportion of his wealth as set by the bank:

$$b = (\lambda - 1)z, \quad \lambda > 1.$$

AB used U.S. data and a two stage estimation procedure to test the exogeneity of the capital constraint (under the null hypothesis of EJ, it is exogenous). They include measures of both transferable human capital (education, etc.) and of entrepreneurial ability (business experience, etc.). At the first stage they examined the relationship between an owner's human capital, entrepreneurial ability and financial wealth; at the second stage they examine the relationship between the firm's start-up capital, entrepreneurial ability, human capital and financial wealth, with financial wealth the *predicted* value determined from the first step.<sup>21</sup>

AB found that at the first stage wealth increased with both human capital and entrepreneurial ability ( $z = z(x, \theta)$ ,  $z_x, z_\theta > 0$  is demonstrated), suggesting that collateral constraints are *endogenous*, contrary to EJ's assumption. At the second stage start-up capital was found to increase with entrepreneurial ability and human capital controlling for financial wealth ( $k = k(z, x, \theta)$ ,  $k_\theta > 0$  is demonstrated), suggesting that better-quality entrepreneurs *at a given level of human capital* are more credit constrained, consistent with the EJ model.<sup>22</sup> Interestingly, the marginal effect of wealth was diminished significantly once *human capital* is added ( $\partial^2 k / \partial z \partial x < 0$ ), whereas adding *entrepreneurial ability* increased the marginal effect of wealth ( $\partial^2 k / \partial z \partial \theta > 0$ ). Thus human capital *mitigates* wealth constraints which themselves had a greater impact on better entrepreneurs. Capital constraints are thus (as Cressy, 1996 found) endogenous, but importantly, controlling for human capital and entrepreneurial ability *does not completely eliminate them*. Astebro and Bernhardt were also able to control for the fact that some industries have a larger efficient minimum scale (MES—see Chapter 7 in this volume) and for the fact that there are differences in risk across industries, both of which may militate against the decision to enter. Industries with larger MES are entered by wealthier entrepreneurs and firms with greater risk of failure start with less capital, suggesting that credit constraints bite more strongly in these categories.

Parker and Van Praag (2004) (henceforth PVP) also addressed the important endogeneity issue using Dutch self-employment data and a model

developed in Bernhardt (2000). They used a different definition of capital constraints, namely, the practice of loan *down-scaling*, where the bank may offer less than the firm requests, rather than assuming (as in EJ, etc.) that the bank offers to lend in proportion to the firm's assets. PVP argued that downscaling of a loan is evidence for the existence of a credit constraint on a firm. This is plausible if one were able to control adequately for other factors that might explain the downscaling. These include the degree of optimism of the borrower (de Meza and Southey, 1996). Since younger borrowers are more likely to have their applications downscaled (as optimists they will ask for too much), PVP's entrepreneurial age variable in effect controls for this possibility.

PVP's study demonstrated empirically the existence of an "endogenous triangle" relating human capital, capital constraints and performance among the Dutch self-employed. They found (confirming the AB result) above that credit constraints measured by the extent of loan downscaling are endogenous, being lower for individuals with greater human capital. Thus more educated individuals are less restricted in starting a business because they are better capitalized (i.e., possess more initial assets). Capital constraints in PVP in turn were found to impede performance (measured by profits from the business) since they constrain it to a suboptimal initial scale. Finally, human capital enhances business performance directly (via entrepreneurial ability's effect on productivity) and indirectly (via a relaxation of capital constraints). The effects of human capital on constraint relaxation are also, they argue, *quantitatively* important, suggesting a role for public policy in promoting entrepreneurship education at school (see Chapter 4 of this volume). The total rate of return to schooling for entrepreneurs is of the order of 13.5–15%. This is clearly a serious consideration when designing overall policies directed at promoting business activity.<sup>23</sup>

*5.2.3. The Role of Savings* Another deficiency of the basic EJ model is that it fails to allow for the fact that an individual may save and use this to finance his business activity in the next period (Xu, 1998). Testing for this, however, (see the discussion of AB above) has been shown not to eliminate the correlation between assets and the switching decision.

*5.2.4. Sample Issues* Other criticisms of EJ revolve around the dataset EJ and others (Holtz-Eakin et al., 1994b) use. EJ's original sample was of American young white males with an average wealth of \$20,000, a very modest figure indeed. Cressy (1996) also worked with low wealth individuals. In contrast, the sample used in a follow-up study by Holtz-Eakin et al. (1994a) was of rather richer U.S. individuals with an average wealth of \$73,000 in 1981.<sup>24</sup> While the original Holtz-Eakin et al. article suggested the existence of capital constrained entrepreneurs *throughout* the wealth spectrum, a recent paper by

Hurst and Lusardi (2004) examining a wider sample of U.S. citizens suggested that throughout most of the wealth range there is in fact *no* correlation between the chances of starting a business and individual wealth levels. They found that it is only at the very top of the distribution that the correlation becomes positive. Thus only very high net worth individuals have “excessive” entrepreneurial ability relative to their assets and are thereby credit constrained. Thus studies such as Holz-Eakin et al., that used samples of high net worth individuals to argue more generally for the existence of credit constraints seem now to be much more restrictive in scope than had been previously envisaged.

The policy implications derived from this finding are, however, inherently implausible or at least seemingly unjust. Government interest in intervening in markets with credit constraints has traditionally focused on employment issues, namely, whether relatively poor, often unemployed, low ability, cash-strapped individuals can efficiently start their own firms. It now appears from Hurst and Lusardi’s results that only the *richest* individuals in society are “cash-strapped:” individuals who are likely to have plenty of talent and to be the least likely to be unemployed presenting socially concerned policy-makers with a rather exquisite paradox! It may, of course, *logically* be the case that such individuals are indeed the ones to target with loan guarantee schemes and government subsidies because in the long run we shall reap the benefits in the form of higher national wealth. However, for most right-minded economists adopting the motto “to him who hath it shall be given” is likely to stick in the throat.

*5.2.5. Contrarian Evidence* Other studies of capital constraints from different methodological perspectives to the EJ tradition have tended to conclude that credit constraints are in general of little importance. For example, Aston Business School (1991) in a survey of potentially fast growth businesses found that at most 6% of businesses with growth potential were constrained. This is a rather small proportion of businesses if capital constraints are as widespread as the empirical work suggests. Likewise, Cambridge’s Business Research Centre (1998) and the Bank of England (1996), also adopting a survey approach, found little evidence of financial constraints among U.K. high-technology small businesses.

### *5.3. Alternative Explanations for the Findings*

The fact that a theory is consistent with the data does not, of course, prove that it explains it. Thus in the following sections we examine alternative plausible explanations for the correlation of assets and entrepreneurial switching/survival.

*5.3.1. Risk Aversion of the Entrepreneur* Another potential explanation exists for the EJ finding that assets and switching into SE are positively correlated but which does not imply the existence of credit constraints. This explanation depends merely on some plausible assumptions and limited evidence about human tolerance of risk. It is commonly believed (and there is evidence to show) that people in general *dislike* risk. Studies of the stock market show that people need to be offered higher returns to invest in more risky securities. This is consistent with risk aversion. Likewise, most people take out some kind of insurance policy against fire, theft, etc., which involves the payment of a premium. This also suggests dislike of risk since by the mechanism of insurance the risk is transferred to another party.<sup>25</sup>

Imagine, then, that when I increase your assets you become less risk averse, that is, you become more willing to take risks: for example, if I offer you simultaneously an increase in your wealth  $W$  by £1 and a bet which yields +£1 with probability 1/2 and -£1 with probability 1/2, *with your additional assets* you are now *more* likely to take the bet than before. In the language of economics this means your utility of income function displays Decreasing Absolute Risk Aversion (or DARA). Since available empirical evidence suggests that entrepreneurship tends to be more income-risky than wage employment, this means that the marginal entrepreneur (one for whom the expected costs *just outweigh* the expected benefits) would switch into self-employment should he or she receive a windfall gain. There is, furthermore, some empirical evidence to support the assumption that entrepreneurs have decreasing absolute risk aversion (see Guiso and Paiella, 1999).

Thus we have the result that higher wealth is associated with greater propensity to enter SE which gives us the EJ result but without capital constraints. No direct test of this proposition is yet available even though it would be straightforward to set up.

*5.3.2. Control Aversion of the Entrepreneur* Entrepreneurs of smaller firms are well known to be *control averse*.<sup>26</sup> Control aversion is defined here as the dislike of perceived interference by outsiders in a business. Control aversion among small firms may in general affect their decision to take on external equity providers or their decision to take on external debt. Empirically there is growing evidence that such aversion does exist and influences both the capital structure and performance of small firms (Cressy and Olofsson, 1997; Mueller, 2004). So how can it explain the empirical results on credit constraints?

Entrepreneurs do not like any kind of interference in their operations, in particular by the local bank manager (Cressy, 1995). Partly for this reason, and as noted earlier, they tend to borrow relatively little. In the language of economics this means that (at the margin) the psychological costs of borrowing at the profit maximizing level outweigh the benefits. As firms get larger things

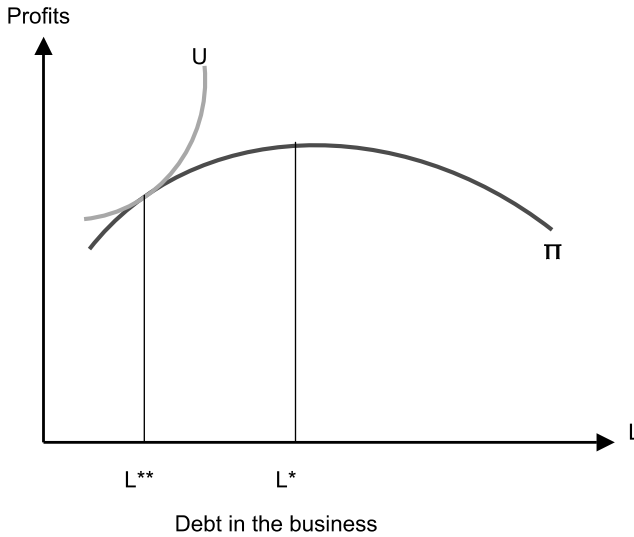


FIGURE 8-4 *Effects of control aversion on the amount of borrowing.*

get less personal, management tends to be rewarded by salaries rather than simply profits, and the aversion to perceived bank interference starts to wane. But at the level of the micro business (one with fewer than ten employees) control-aversion is likely to restrict borrowing not from the supply *but from the demand side*.

The equilibrium tradeoff is illustrated in Figure 8-4 (taken from Cressy, 1995) where the red line indicates profits of the firm as a function of borrowing. This represents the utility function of the financial manager of a larger firm. By contrast the green line represents an indifference curve for the entrepreneur of a small firm. While profits are a “good” (yield positive marginal utility) to such an individual borrowing is “bad” (yields negative marginal utility). Thus the indifference curve is upward-sloping—its slope being the ratio of the marginal utility of borrowing to that of profits. Utility is therefore increasing as we move to the North-West of the diagram with higher profits and lower borrowing. The highest indifference curve attainable with the red profit constraint is the green one. The optimum for the larger firm is where profits are maximized, at  $L^*$ . The optimum of the control-averse entrepreneur equates the marginal disutility of borrowing with the marginal utility of profits, yielding the smaller borrowing amount  $L^{**}$ .

## 6. RATIONING, BANK HETEROGENEITY AND BORROWER SEARCH

One feature of the definition of credit rationing is that we assume that the potential borrower has approached all potential lending sources and discovered that none will lend any money. In the simplest competitive model, of course, all banks are identical, potential borrowers are costlessly aware of offers of funds and therefore search becomes unnecessary. However, we can imagine a world (the real world!) in which banks are heterogeneous, information is costly to the entrepreneur to acquire and that as a result a borrower approaches only a subset of potential lenders. (Indeed, some empirical research shows that it is highly likely that banks do *not* all have the same beliefs about a given project.) This might arise, for example, if manager ability (experience) is in short supply and distributed unevenly across banks. Then some banks will be more informed than others. This may in turn lead to one bank rejecting the firm's request and another to accepting it. But once again, if the supply of credit from any bank is elastic, and search is costless, entrepreneurs will find the banks that are more informed about their profitable project and borrow from them, so the banks that are less informed will have losses.<sup>27</sup> It is only if firms do not know the 'best' bank and there are costs to finding this out that they may end up starved of funds for good projects. But this does not have to be a market imperfection: if at the margin the costs of search are equated to the expected benefits, then just the right number of projects will be funded.

## 7. MARKET SOLUTIONS

Before crying wolf about credit constraints or credit rationing it is important to see how, if at all, the market is already responding to the perceived problem. In this section we briefly examine the role of relationship banking, past and present, as a solution to the potential rationing of credit. This is followed by subsections on mutual guarantee schemes and outside equity provision.

### *7.1. Relationship Banking*

Relationship banking describes how a bank manager through his/her long term personal relationship with the entrepreneur(s) acquires information about the firm that will be material to his/her lending decisions (amounts, prices, collateral, etc.). Empirically, small firms rarely switch banks—the cost of doing so is too high. And the result is that the typical firm has a relationship of some ten years with its bank (Berger and Udell, 2002). Relationship banking would seem advantageous to a high-quality firm which expects better credit facilities and perhaps less rationing than otherwise identical firms with which



the bank has a shorter relationship. By contrast, a low-quality firm might expect progressively less favorable treatment over time as the bank learns its deficiencies. In practice, the evidence demonstrates that relationship banking results in a lower requirement for collateral (Berger and Udell, 1995) and more credit availability, but seems to have little affect on the interest rates charged (Petersen and Rajan, 1994).

Empirical evidence from the U.S. (Petersen and Rajan, 1994)—henceforth PR—shows that small firms are (as in the U.K.) highly dependent on a single lender. PR report that in the U.S. the smallest firms (measured by assets) borrow 95% of their funds from the largest lender available to them. This falls to “only” 76% for the largest firms in the sample. Cressy (1993) finds a similar pattern for the U.K. with start-ups generally borrowing (if at all) from one bank and less than 10% from other banks as well. He also finds that borrowers with fewer relationships with other banks tend to pay higher interest rates. There may be several reasons for this. For example, if collateral is being spread more thinly by the multiple-relationship firm, we should expect interest rates charged by its main bank to be higher to cover the higher expected loss from a given size loan. Likewise, if the main bank has little control over the firm’s borrowing from other sources it may consider the current loan more risky (the true debt-equity ratio is unknown) and charge a higher interest rate to compensate for that extra risk.

PR argue that relationship lending should lower the interest rate paid by a borrower as asymmetric information and costly adverse selection are removed. Interestingly, however, the data contradict this: they find no empirically significant effect of relationship duration on the average interest rate paid. If, however, we take the view that the bank learns about the firm’s quality over time and prices accordingly, we should not expect the *average* interest rate necessarily to decline, since borrowers will be revealed to be either good or bad. What we should in fact find is that the *dispersion* of interest rates should *increase* over time as the sheep (good borrowers) are separated from the goats (bad borrowers) and priced accordingly. We should also find that credit rationing decreases over time. Credit rationing occurs in a situation in which the loan supply curve of the bank is backward bending: as the price of funds is increased, the quality of borrowers declines making further expansion of supply unprofitable to the bank and creating an equilibrium in which the supply of funds is permanently lower than the demand at *all* interest rates (see Parker, 2002). This situation results in the phenomenon of rationing where some borrowers are (randomly) denied funds for viable projects. Relationship banking, in (gradually) making borrower quality observable, then allows the bank to price quality differentially thus avoiding the rationing effects of adverse selection. So, operationalizing this, PR predict that relationship banking should increase the availability of credit to smaller businesses. Examining the availability of credit empirically

they find that longer relationships between U.S. small firms and their banks are indeed associated with greater credit availability to those firms.

Recent evidence on the effects of major changes in the U.S. banking system suggests a decline in the importance of relationship banking over the period 1973–1997 (Petersen and Rajan, 2002). Borrowers have become physically more distant and communicate progressively less in person with their bank. PR provide evidence that this is due not to branch closures, more lax credit standards (predicting increasing bankruptcies) or changes in the nature of the firms borrowing (e.g., gravitating to rural rather than urban locations). They show that the changes are due to increases in bank productivity as a result of advances in IT and information collection, allowing the soft (i.e., qualitative) information generated by relationships to be progressively replaced by hard (i.e., quantitative) information from other sources. This extra hard information includes such things as whether a borrower is current in his trade credit payments, to whom he has applied for credit, and so on. Firms that borrow at a distance, they argue, can now be of *lower* credit quality because the new technology of information gathering allows lenders to make loans to them with less fear of loss. This view is buttressed by the evidence which shows (a) that distance is indeed a good predictor of borrower quality and (b) distance is becoming less useful as a predictor over time.

## 7.2. Mutual Guarantee Schemes

A mutual guarantee scheme (MGS) is an arrangement whereby firms in a given locality or industry pay a membership fee to join an association for (among other things) insurance against loan default. The arrangement is that when any member firm approaches a bank to borrow, the MGS will guarantee to repay the loan if the firm defaults. The chances of the MGS not being able to pay the bank are less than that of any individual firm (though the precise probability that an MGS itself will default depends on how large its membership is and the upper limit to borrowing by any given member). This gives the bank an incentive to lend to firms with inadequate collateral. Likewise, if the cost of membership is not too high relative to the benefits of borrowing, there is an incentive for the firm to join the MGS. MGS societies also have a strong incentive to monitor potential and actual members to avoid unscrupulous behavior and scrupulous risk as both these things may result in MGS default if prevalent enough. However, such monitoring is feasible only within a reasonably well-defined and tightly knit group of firms, for example, in a given locality and industry or craft, meaning that the scope of MGS is limited. Furthermore, these “collective” approaches to borrowing may be more successful in some countries than others due to cultural reasons. For example, it appears that the Southern countries of the EU together with France (which straddles Northern

and Southern regions) are more amenable to MGS societies than the Northern ones. It appears that the countries of Northern Europe, where collective action is less of a tradition, are less willing hosts to this kind of organization. This is another reason why mutual guarantee schemes offer limited scope for dealing with widespread credit constraints.

### 7.3. *Outside Equity*

If debt cannot easily be raised by a small firm due to absence of collateral then one might imagine that equity would be the alternative and indeed more suitable form of finance. Outside equity funding involves the purchase by an outside organization or individual of shares in the firm. It requires no fixed payments schedule from the user and no collateral is required to support its provision. However, despite its seeming attractiveness there are insurmountable problems to its use in addressing alleged credit constraints on small businesses.

First, outside equity is by definition irrelevant to the majority of small businesses that are unincorporated and hence cannot (legally) issue equity. The formation of partnerships may offer some kind of solution to sole traders facing credit constraints, although the greater degree of risk exposure implied by the increased liability of such partnerships may often dissuade sole traders from choosing this option. Second, even if we confine our interest to small incorporated businesses, control aversion operates even more strongly in the case of equity (by comparison with debt) to discourage most small firms from gaining finance this way.<sup>28</sup> The decision not to take advantage of outside equity may well result in more gearing and slower growth for the firms involved, but their owners seem to prefer the disutility of slower growth to the disutility of control-loss (Mueller, 2004).<sup>29</sup> Third, venture capitalists or business angels—the likely source of such finance—are not interested in buying equity in the vast majority of small limited companies (*typical* firms) as they offer no prospects of capital gain of the order they are used to and require. Traditionally VCs claim they have rates of return (IRRs) in excess of 30% per annum on their investments (although these figures may be biased upward—see Cochran, 2002). Such rates of return are only possible however if the firm grows very fast and in a short time (3–5 years) ends up with a stock market flotation or a trade sale. The vast majority of firms (at a rough guess well over 95%), and even the majority of *sophisticated* firms, do not fall into this category. Bhide (1999), for example, found that the vast majority of his American fast growth firms grew from retained profits rather than venture capital.

In summary, outside equity does not promise to alleviate widespread credit constraints should they exist as a result of both supply- and demand-side issues in the financial markets. Such issues do not in general constitute a

market imperfection, however, and therefore do not automatically constitute an invitation to government intervention.

## 8. GOVERNMENT SOLUTIONS

Government can intervene in the credit market in many different ways in response to perceived shortages of funds for small firms. One of the more popular and arguably more successful methods of intervention is via loan guarantees.

### 8.1. Loan Guarantee Schemes

If relationship banking only works within a time horizon of five years or more and MGS schemes require specific local and cultural conditions to solve credit shortages, we might conclude that government intervention is desirable at least for the well-defined subset of young, small firms under consideration. And one of the most popular remedies proposed by governments is the Loan Guarantee Scheme or LGS. But we shall see that these schemes are by no means bereft of theoretical difficulties, whatever their practical usefulness may turn out to be.

Under a loan guarantee scheme, the government agrees to indemnify the bank up to a certain proportion of its loan to a borrower without collateral, with the interest rate charged “on purely commercial grounds,” in return for the borrower paying an insurance premium on the loan of 1 or 2% of its value. The objective is to get the bank to lend to borrowers to whom it would not otherwise lend in view of the borrower’s lack of collateral. The theoretical problem with such schemes is that they do not address the issue of adverse selection and moral hazard on which they were predicated. Adverse selection, as we have seen above, under conditions of unobservable borrower quality (talent), is dealt with by the bank by making better quality borrowers provide more collateral. Moral hazard, in the context of unobservable borrower effort, is dealt with by the bank by asking for collateral from borrowers. This, as we have seen above, creates an incentive among lazy borrowers to put in effort since, by so doing, they are more likely to avoid losing their house!

Unfortunately, neither of these issues is dealt with by the government loan guarantee scheme since by definition the borrower has not been required to place collateral on her loan, and so the bank is unable to charge the good quality applicant differentially from the bad. Likewise, greater effort cannot be engineered by the LGS for the same reason. Of course it is true that some loans are made without collateral; but *ex hypothesi* this is not the case here.

This might seem to rule out a useful role for a LGS. However, recent empirical research has questioned the relevance of adverse selection in (small) business banking (Toivanen and Cressy, 2002), and while the finding of that research is that moral hazard is very much still with us, the need to use collateral as a sorting device is perhaps no longer quite so compelling. Thus we may need only provide the incentive to effort otherwise missing in the LGS scheme. Furthermore, several rather competent studies in the U.K. suggest that LGS performs better than might have been predicted from theoretical considerations alone. Thus several studies financed by the Department of Trade and Industry (e.g., National Economic Research Associates, 1989) and independent studies such as Cowling and Mitchell (2003) have shown rather convincingly that there is financial and economic additivity,<sup>30</sup> at modest levels, in the workings of the LGS. Thus, firms are getting money they would not have otherwise received and producing output that other firms would not have produced in the absence of the scheme.

## 9. CONCLUSIONS

In this chapter we have examined the nature and importance of bank lending to small firms, the constraints that apparently prevent small firms getting the right amounts of money and the range of solutions offered by the market and government to deal with shortages resulting from credit market imperfections. We have also noted the potentially important role of control aversion and limited objectives of the entrepreneur in the credit demand equation that has only recently begun to be quantified. Thus it is timely to ask ourselves: Are there credit constraints on small businesses in the real world? Is credit rationing an important phenomenon?

These questions can, of course, only be answered, the above discussion suggests, by a detailed *empirical* examination of the characteristics of firms, the sectors of the economy in which they operate, the specific the time period or part of the macro cycle in question and the nature of the information regime in which all this is embedded. Just as one should always consult more than one doctor for a health check, no one theory should be relied upon without a second opinion. In my view, the empirical results that seem to me convincing rely less on theory for justification than straightforward questioning of participants together with cross-checks from other sources. By and large, these kinds of studies suggest that credit constraints are not a widespread phenomenon and that effective government intervention in the small minority of cases where it may exist is cheap and effective. Therefore, despite the mountain of theoretical literature suggesting the abstract possibility of credit constraints, it does not appear in general to be an important empirical phenomenon (Cressy, 2002; Parker, 2002).

Often the major obstruction to small firm development in a country is not credit rationing as such but rather the *competitiveness of the banking system* and *the amount of bank information about small business profitability*. Some personal experience is instructive here. At a recent conference held in Zagreb, Croatia, which focused on the role of loan and mutual guarantee schemes in former Eastern Bloc countries, it became clear that in many of these countries there was a dominant oligopoly in the banking system which preferred to concentrate on large international firms rather than small local ones. The banks therefore had little awareness of the potential market for loans from their entrepreneurial firms—one long exploited by banks in the West. Part of the problem was clearly the issue of the availability of collateral where loan or mutual guarantee schemes might be useful, but the main driver seemed to be the absence of knowledge of small firms and their financial requirements by the big banks. Banking reform may therefore be the first priority and without it other changes would be simply “rearrange the deckchairs on the Titanic.”

## NOTES

<sup>1</sup> Thus we shall spend some time in this chapter discussing the empirical tests of the Evans and Jovanovic (1989) theory but shall not enter at all into a discussion of the literature emanating from Fazzari, Hubbard and Petersen’s (1987) study of the effect of cash flow on investment. The latter is excluded as it is not focused specifically on bank lending issues but rather on the availability of finance in general.

<sup>2</sup> Bhidé’s (1999) concept of the *promising* business corresponds rather closely to what we call the sophisticated small business.

<sup>3</sup> A defining characteristic of the sophisticated small firm described in Bhidé (1999) is that it is one of the (100) fastest growing firms in the (U.S.) economy where growth is measured by growth of sales.

<sup>4</sup> Cressy shows that the proportion borrowing rises to 50% after three years, with borrowing concentrated on the higher quality businesses. The figures are partly explained by the fact that the typical business start-up is a very paltry affair, often not owning or even leasing premises and employing few people other than the owner and perhaps his/her spouse.

<sup>5</sup> Small is defined as fewer than 500 employees. The median book value of assets was \$130,000 and median annual sales \$300,000. The median age of firm was ten years.

<sup>6</sup> This appears paradoxical but the two propositions are mutually consistent if one allows for the fact that the probability of borrowing from non-bank sources in the States is decreasing at a faster rate than overall borrowing.

<sup>7</sup> In the U.S., the term “commitment lending” refers to overdrafts or lines of credit since the bank is committing itself to maintaining a fixed margin regardless of the Base or Prime rate prevailing over the next three months.

<sup>8</sup> Thus we are ignoring the potential mismatch of payments for purchases and sales, another motive for overdrafts.

<sup>9</sup> We assume again for simplicity that no interest is paid on this overdraft. This adds more mathematics but no more clarity to the basic idea.

<sup>10</sup> Evaluated at the sample mean. The relationship is in fact inverse U-shaped peaking at around 55 years of age.

<sup>11</sup> We follow accounting practice and define the long term as a period greater than one year.

<sup>12</sup> One Swedish bank manager is quoted as saying that the problem with service firms is that the assets are their people, and they tend to go home at 5 o'clock! See Cressy and Olofsson (1997).

<sup>13</sup> A more specialized asset is one which is inherently more difficult to convert into cash. Many intangible assets fall into this category, for example, patents and copyrights.

<sup>14</sup> The reasons for this are almost certainly (a) most loans were taken out before the recession started (approximately in 1990) and (b) most defaults occur early in the term.

<sup>15</sup> It is well known that both larger and older firms are less failure/closure prone. See, for example, Evans (1987).

<sup>16</sup> Borrower effort is not observed in practice because it is too costly in terms of bank manager time. The typical bank manager in the U.K. will have in the region of 200 small business customers to deal with. Visiting their premises is a very costly business relative to the advantages it offers.

<sup>17</sup> See Cressy (2002) for a summary of a recent debate by experts in this area.

<sup>18</sup> We have seen that lending rules are not quite as simple as that in practice which makes the theory questionable from the outset. In particular, even if the approximation is good for start-ups (the EJ dataset) it is not necessarily good for larger or more established smaller businesses. There is also a question about the individual wealth levels to which the (the basis of collateral) theory applies. We shall discuss this issue later.

<sup>19</sup> The actual estimation sample is 1443 since negative net worth or SE income individuals were deleted from the sample. This is, of course, a potential source of bias.

<sup>20</sup> This model is taken from Astebro and Bernhardt (2003). However, there is a more recent paper by Parker and van Praag (2004) that tests similar ideas in a slightly different theoretical framework proposed by Bernhardt (2000).

<sup>21</sup> Identification of this instrument is accomplished by using county-level indicators of household income for the owners.

<sup>22</sup> Their model predicts that individuals with greater entrepreneurial ability for *given* wealth will be more credit constrained.

<sup>23</sup> This appears to be an average rather than a marginal return, however, and so may overstate the return at the margin.

<sup>24</sup> Holtz-Eakin et al. following Blanchflower and Oswald (1998) examined the impact of both the wealth of the individual defined as one's liquid assets and house equity and any inheritance on the decision to enter business and the capitalization of the business once started. In fact, only the inheritance variable had any impact and that impact was quite substantial. For example, a \$100,000 inheritance increased the probability of transition into SE by about 15%. Importantly, Holtz-Eakin et al. also showed that the inheritance effect is not due to the inheritance of businesses. If the latter were true then the observed correlation of inheritance and start-up propensity would simply have been the decision by inheritors to continue running their parents' businesses.

<sup>25</sup> There are, of course, counterexamples. The most glaring is the fact that huge numbers (millions) of people, often the poorest, engage in regularly in an unfair bet, namely the national lottery. This is inconsistent with risk aversion.

<sup>26</sup> Evidence for this goes back at least 30 years to the U.K.'s Bolton Committee a landmark in the study of smaller firms (Bolton, 1971). However, of more recent vintage and referring specifically to aversion to bank control is Cressy (1995).

<sup>27</sup> We can imagine that the anticipation of such losses would lead to banks competing up the price of scarce informed managers by their attempting to attract them away, and so on.

<sup>28</sup> Cressy and Olofsson (1997) found that some small Swedish firms would rather sell the business altogether than give up a share to an outsider! The aversion to outside equity was to some degree lower for younger firms and firms in the service industries.

<sup>29</sup> A greater degree of financial underdiversification on the part of the entrepreneur is also a consequence of not taking on outside equity.

<sup>30</sup> Financial additivity occurs if the funds provided by LGS would not have been provided by other private sector financial institutions. Economic additivity occurs if the output from projects financed by LGS does not "crowd out" private sector output.

## REFERENCES

- Astebro, T. and Bernhardt, E. (2003). The Winner's Curse of Human Capital. *Small Business Economics* 24: 1–16.
- Aston Business School (1991). *Constraints on the Growth of Small Firms*. Report for the Department of Trade and Industry, London.
- Bank of England (2003). *Quarterly Report on Small Business Statistics*. Bank of England, London.
- Bank of England (1996). *The Financing of Technology Based Firms*. Bank of England, London.
- Basu, A. and Parker, S. C. (2001). Family Finance and New Business Start-ups. *Oxford Bulletin of Economics & Statistics* 63: 333–358.
- Bates, T. (1990). Entrepreneur Human Capital Inputs and Small Business Longevity. *Review of Economics and Statistics* 72: 551–559.
- Berger, A. and Udell, G. (1990). Collateral, Loan Quality and Bank Risk. *Journal of Monetary Economics* 25: 21–42.
- Berger, A. N. and Udell, G. F. (1992). Some Evidence on the Empirical Significance of Credit Rationing. *Journal of Political Economy* 100: 1046–1077.
- Berger, A. and Udell, G. (1995). Relationship Lending and Lines of Credit in Small Firm Finance. *Journal of Business* 68: 351–381.
- Berger, A. and Udell, G. (1998). The Economics of Small Business Finance: The Roles of Private Equity and Debt Markets in the Financial Growth Cycle. *Journal of Banking and Finance* 22: 613–673.
- Berger, A. and Udell, G. (2002). Small Business Credit Availability and Relationship Lending: The Importance of Bank Organizational Structure. *Economic Journal* 112: F32–F53.
- Bernhardt, D. (2000). Credit rationing? *American Economic Review* 90: 235–239.
- Bhide, A. (1999). *The Origins and Evolution of Small Businesses*. Oxford University Press, Oxford.
- Blanchflower, D. and Oswald, A. (1998). What Makes and Entrepreneur? *Journal of Labor Economics* 16: 26–60.
- Bolton, J. (1971). *Report of the Committee of Enquiry on Small Firms*. Report of Committee chaired by Sir John Bolton. HSMO, London.
- Boot, A., Thakor, A. V. and Udell, G. F. (1991). Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results. *Economic Journal* 101: 458–472.



- Brealey, R. and Myers, S. (2003). *Principles of Corporate Finance* (7<sup>th</sup> edition). McGraw-Hill, New York.
- Business Research Centre (1998). *The State of British Enterprise: Growth, Innovation and Competitive Advantage in Small and Medium-sized Firms*. Mimeo, Department of Economics, University of Cambridge.
- Chan, Y. and Thakor, A. V. (1987). Collateral and Competitive Equilibria with Moral Hazard and Private Information. *Journal of Finance* 42: 345–363.
- Coco, G. (2000). On the Use of Collateral. *Journal of Economic Surveys* 14: 191–214.
- Cosh, A. and Hughes, A. (1994). Size, Financial Structure and Profitability. In *Financing Small Firms*, A. Hughes and D. J. Storey (Eds). Routledge, London.
- Cowling, M. and Mitchell, P. (1996). *A Review of the Loan Guarantee Scheme*. Occasional Paper, Warwick Business School, University of Warwick, England.
- Cowling, M. and Mitchell, P. (2003). Is the Small Firms Loan Guarantee Scheme Hazardous for Banks or Helpful to Small Businesses? *Small Business Economics* 21: 63–72.
- Cressy, R. C. (1993). *The Start-up Tracking Exercise: Third Year Report*. Report prepared for National Westminster Bank of Great Britain.
- Cressy, R. C. (1995). Borrowing and Control: A Theory of Business Types. *Small Business Economics* 7: 1–10.
- Cressy, R. (1996). Are Business Start-ups Debt-Rationed? *Economic Journal* 106: 1253–1270.
- Cressy, R. (1998). Credit Rationing or Risk Aversion? An Alternative Explanation for the Evans-Jovanovic Finding. *Economics Letters* 66: 235–240.
- Cressy, R. (2002). Introduction: Funding Gaps. *Economic Journal* 112: F1–F16.
- Cressy, R. (2004). Why do Most Firms Die Young? *Small Business Economics* (forthcoming).
- Cressy, R. and Olofsson, C. (1997). European SME Financing: An Overview. *Small Business Economics* 9: 87–96.
- Cressy, R. and Storey, D. (1994). *New Firms and Their Bank*. Report, National Westminster Bank, Great Britain.
- Cressy, R. and Toivanen, O. (2001). Is There Adverse Selection in the Credit Market? *Venture Capital* 3: 215–238.
- Damodaran, A. (1999). *Applied Corporate Finance: A User's Manual*. Wiley, New York.
- De Meza, D. and Webb, D. (1987). Too Much Investment: A Problem of Asymmetric Information. *Quarterly Journal of Economics* 102: 281–292.
- ENSR (1997). *The European Observatory for SMEs, Fifth Annual Report to the European Commission*. EIM, Netherlands.
- Evans, D. S. (1987). The Relationship Between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries. *Journal of Industrial Economics* 35: 567–581.
- Evans, D. and Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints. *Journal of Political Economy* 97: 808–827.
- Evans, D. and Leighton, L. (1989). Some Empirical Aspects of Entrepreneurship. *American Economic Review* 79: 519–535.
- Fazzari, S. M., Hubbard, R. G. and Peterson, B. (1988). Financing Constraints and Corporate Investment. *Brookings Papers on Economic Activity* 1: 141–195.
- Guiso, L. and Paiella, M. (1999). *Risk Aversion, Wealth and Background Risk*. Unpublished Manuscript, Birkbeck College, London.
- Hamilton, B. H. (2000). Does Entrepreneurship Pay? *Journal of Political Economy* 108: 604–631.
- Han, L. (2005). *Collateralization, Interest Rates and Signaling in Entrepreneurial Finance*. PhD Dissertation, Warwick Business School, University of Warwick, England.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. S. (1994a). Sticking It Out: Entrepreneurial Survival and Liquidity Constraints. *Journal of Political Economy* 102: 53–75.

- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. S. (1994b). Entrepreneurial Decisions and Liquidity Constraints. *RAND Journal of Economics* 25: 34–347.
- Hughes, A. and Storey, D. (1994). *Financing Small Firms*. Routledge, London.
- Hurst, E. and Lusardi, A. (2004). Liquidity Constraints, Household Wealth and Entrepreneurship. *Journal of Political Economy* 112: 319–347.
- Leeth, J. D. and Scott, J. A. (1989). The Incidence of Secured Debt: Evidence from the Small Business Community. *Journal of Financial and Quantitative Analysis* 24: 379–393.
- Mueller, E. (2004). *The Performance of Private Companies: An Empirical Investigation into the Role of Control, Risk and Incentives*. PhD Dissertation, London School of Economics, London.
- National Economic Research Associates (NERA) (1989). *An Evaluation of the Loan Guarantee Scheme*. Research Paper No. 74, Department of Employment, London.
- Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Parker, S. C. (2002). Do Banks Ration Credit to New Enterprises? And Should Governments Intervene? *Scottish Journal of Political Economy* 49: 162–195.
- Parker, S. C. and Van Praag, C. M. (2004). *Schooling, Capital Constraints and Entrepreneurial Performance: The Endogenous Triangle*. Working Paper 45/04, Tinbergen Institute, University of Amsterdam, <http://www1.fee.uva.nl/scholar/>.
- Peterson, M. A. and Rajan, R. G. (1994). The Benefits of Lending Relationships: Evidence From Small Business Data. *Journal of Finance* 49: 3–37.
- Petersen, M. A. and Rajan, R. (1995). The Effect of Credit Market Competition on Lending Relationships. *Quarterly Journal of Economics* 110: 407–443.
- Petersen, M. A. and Rajan, R. (2002). Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance* 57: 2533–2569.
- Shamandoura, T. (2003). *Credit Risk of Loan Contracts: An Application to Small and Medium Businesses in the U.K.* PhD Dissertation, Department of Economics, University of Essex.
- Stiglitz, J. E. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review* 71: 393–410.
- Storey, D., Keasey, K., Watson, R. and Wynarczyk, P. (1987). *The Performance of Small Firms*. Croom Helm, Beckenham, Kent.
- Treasury, H. M. (2000). *Competition in U.K Banking: A Report to the Chancellor of the Exchequer*. Report by Donald Cruickshank, London: [http://www.hmTreasury.gov.uk/documents/financial\\_services/banking/bankreview/fin\\_bank\\_reviewfinal.cfm](http://www.hmTreasury.gov.uk/documents/financial_services/banking/bankreview/fin_bank_reviewfinal.cfm).
- Watson, R. (1990). Employment Change, Profits and Directors' Remuneration in Small and Closely-held U.K Companies. *Scottish Journal of Political Economy* 37: 259–274.
- Xu, B. (1998). A Re-Estimation of the Evans-Jovanovic Entrepreneurial Choice Model. *Economics Letters* 58: 91–95.

CHRISTIAN KEUSCHNIGG<sup>a</sup> and SØREN BO NIELSEN<sup>b</sup>

<sup>a</sup>*University of St Gallen, Varnbuelstrasse 19, CH-9000 St. Gallen, Switzerland  
E-mail: christian.keuschnigg@unisg.ch*

<sup>b</sup>*Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark  
E-mail: sbn.eco@cbs.dk*

## 9. Public Policy, Start-up Entrepreneurship and the Market for Venture Capital

### 1. INTRODUCTION

Among new entrepreneurial firms in high-technology industries, venture capital (VC) has increasingly become an important player, not only as a source of finance but also as a source of professional support. The firm's transition from birth of the idea to a marketable and profitable product not only involves technological experiments and development of prototypes. Acquiring new facilities, developing marketing strategies, attracting key clients and reliable suppliers, hiring new personnel, team building and raising further financing to expand the business require formidable managerial expertise and entrepreneurial experience. While proficient at the technological side, start-up entrepreneurs not only lack the necessary capital but are typically also in dire need of professional assistance. Seasoned venture capitalists (VCs) are well suited to fill these gaps. They have good access to capital, are endowed with own managerial experience and detailed knowledge of the industry. They can count on a well developed network of suppliers, customers and key personnel. Indeed, the defining characteristic of VC is the combination of finance and commercial assistance. In contrast to passive bank financing, VCs arrange for entrepreneurs to receive support in various ways by creating links to suppliers and possible customers, getting hold of key personnel, providing strategic and marketing advice and helping the professionalization of the firm in other ways.

Venture capital started out in the U.S. half a century ago and has grown vigorously in the last 20 years. Almost half of new firms in the U.S. which are sold off at IPOs (Initial Public Offerings) have been backed by VC

(see Gompers and Lerner, 2001). In Europe, the introduction of VC started significantly later, and only in the most recent years have VC firms become prominent financiers of young technology firms. Recent statistics published by EVCA (the European Private Equity and Venture Capital Association) report a total investment by members of the Association of 37 billion Euros in 2004, up from 5.5 billion in 1995.<sup>1</sup> As in previous years, strict seed and start-up investment constituted a minor part of the total amount (some €2.7 billion or 7.3%). The rest was absorbed by financing buyouts and by expansion-stage investments (26.6 and €7.9 billion, respectively). The EVCA statistics further reveal marked differences across countries. Sweden and The Netherlands had the relatively largest Private Equity/VC markets in Europe (between 0.7 and 0.8% of GDP), followed by the U.K. and France (between 0.4 and 0.6%) while Germany, Denmark and Finland recorded only around 0.2% of GDP.

While VC accounts for only a rather small part of total investment, it tends to be concentrated in the most innovative sectors of the economy. Empirical research for the U.S. by Kortum and Lerner (2000) among others has documented that VC is responsible for a disproportionately large share of overall industrial innovation. According to their results, a dollar of venture capital appears to be about three times more potent in stimulating patents than a dollar of traditional corporate R&D. According to their estimates, VC accounted only for about 3% of corporate R&D from 1983 to 1992 but was responsible for about 8% of industrial innovation in this decade. Given an unchanged potency of venture funding, VC investments should have accounted for about 14% of U.S. innovative activity in 1998. Policy makers and the business community have thus taken a strong interest in healthy conditions for financing new firms and in the development of an active VC industry in particular. Young VC-backed firms are considered an important source of innovation and growth. Several important questions arise when developing a policy perspective. Is there enough risk capital available? Do administrative procedures and requirements hinder entrepreneurship in the first place? Are government grants and subsidies to new firms appropriate? Do taxes block the creation and development of start-ups? Do taxes deter the support and advisory effort of VCs to their portfolio companies?

The VC industry itself certainly considers public policy to be relevant and keeps an eye on whether the general policy environment is suitable to promote the development of private equity and venture capital and to encourage entrepreneurship. For instance, EVCA in 2003 and again in 2004, published a benchmarking report on the tax and legal environment in its member countries (cf. EVCA, 2004).<sup>2</sup> The assessment evaluates 13 indicators relating to both the supply-side (i.e., investors in private equity and VC funds and fund managers investing directly in companies) and the demand-side of private equity and VC (i.e., creation of entrepreneurial firms). Among the tax indicators covered are (i) company tax rates, with special attention to those applicable to small and

medium-sized companies; (ii) capital gains tax rates for individuals; (iii) income tax rates for private individuals; (iv) tax incentives for individual investors investing in private equity; (v) the entrepreneurial environment; and (vi) fiscal incentives to enhance research and development.

The EVCA benchmarking report defines a favorable tax environment by the following criteria: (i) Company tax rates, especially for small and medium sized enterprises should help to support entrepreneurship. (ii) A favorable tax treatment of the sale of unquoted investments in growth companies should strengthen the incentive to entrepreneurial investment. (iii) Income tax rates for private individuals should support, attract and retain human capital, in particular entrepreneurs, researchers and highly qualified company managers. (iv) Tax incentives should be adopted for individual investors investing in private equity funds. (v) Fiscal R&D incentives should be adopted.

The benchmarking report reflects a firm belief that taxes matter for entrepreneurship. Empirical research in public finance indeed testifies to the importance of taxes for entrepreneurship. For example, Rosen (2005), in summarizing his research with a series of co-authors, produces ample evidence that once started, the decisions in new firms regarding employment, capital investment and production are markedly influenced by taxes. Gentry and Hubbard's (2000) empirical analysis demonstrates that the progressivity of the tax schedule is important for entrepreneurship. They argue that the progressivity of the income tax acts like a success tax that taxes successful ventures generating high incomes at particularly high rates and thereby significantly reduces the probability of entrepreneurial entry. Gordon (1998) and Cullen and Gordon (2002), on the other hand, argue that high personal tax rates could actually encourage entrepreneurial activity when individuals are able to exploit the option to incorporate. The argument is that entrepreneurs would choose to be noncorporate in the early stage when the business makes losses. They would then save taxes by offsetting these losses against other personal income. Once the business starts to record profits, an entrepreneur prefers to incorporate in order to exploit low corporate taxes. According to this view, high personal income tax rates can thus encourage entrepreneurship because they imply high tax savings from offsetting losses in the early phase. Boadway and Tremblay (2005) offer a broad overview of the theoretical public finance literature on entrepreneurship and examine various rationales for policy intervention with respect to start-up entrepreneurship.

Apart from this public finance literature on entrepreneurship, there is little theoretical or empirical work on the effects of public policies on VC-financed entrepreneurship. Exceptions are a couple of contributions by Poterba (1989a, b) and Gompers and Lerner (1998). These authors find some evidence of a moderately negative effect of the capital gains tax on VC investments and fund raising. Capital gains taxation tends to depress demand for VC by discouraging

entrepreneurial entry. Since the entrepreneur's income from starting a firm mainly consists of capital gains earned in the start-up period, the capital gains tax makes VC backed entrepreneurship less attractive relative to dependent employment. The capital gains tax can also hamper fundraising since investors' returns mainly consist of capital gains as well. In addition, Gompers and Lerner (1998) found that liberalization of pension fund investment regulations can be an important source of new capital and can thereby stimulate the expansion of the industry.

More recently, Da Rin, Nicodano and Sembenelli (2005) have found that the corporate capital gains tax hurts VC investments in Europe, in particular for early stage investments. None of this empirical literature has actually been able to identify how taxes might change the relative performance of VC-backed compared to other firms by affecting the incentives of VCs to provide support and add value to their portfolio companies. Our own previous theoretical work has aimed to shed light on how exactly taxes as well as subsidies can affect the number of VC-backed firms and the incentives of entrepreneurs and VCs to exploit the full potential of these firms (see Keuschnigg, 2003, 2004a, b; Keuschnigg and Nielsen, 2003a, b, 2004a, b).

The effectiveness of subsidies to capital and research investments of young firms has been investigated empirically by Lerner (1999) and Wallsten (2000), among others. These authors conclude that programs such as the Small Business Innovation Research (SBIR) program in the U.S. can significantly raise the growth of awardee firms compared to other matched firms. This superior performance was confined to awardees in areas with substantial new firm creation. Wallsten found significant crowding out effects although he too argued that the program could help firms to attract additional private funding. The program might thus have a certification role in the sense that participation in the program makes firms more likely to attract additional venture financing. Our theoretical analysis yields an ambiguous conclusion on the desirability of such subsidies. Although successful in boosting the rate of business creation, start-up subsidies may in fact reduce the quality of VC-backed entrepreneurship.

Since Black and Gilson (1998), the presence of specialized stock markets for young technology firms is considered an important precondition for an active venture capital industry. Liquid stock markets allow VCs to exit from their portfolio companies faster and more profitably. This exit possibility also helps the entrepreneur to regain control over the company when the concentrated stake of the VC firm is broadly dispersed over smaller market investors at an IPO. Since entrepreneurial independence is a main motivation for entrepreneurship in the first place, the presence of specialized stock markets makes potential entrepreneurs more willing to start a firm. It also makes the value added of VC financing more attractive to entrepreneurs since the intense control of VCs is expected to last only for a limited time. According to

Micchelacci and Suarez (2004), the presence of liquid stock markets allows VCs to exit faster and to reshuffle their activities to new early stage companies where VC support is needed the most. In equilibrium, liquid stock markets boost innovation and growth because the faster turnover of VC allows for a larger rate of VC backed entrepreneurship. The empirical analysis of Da Rin, Nicodano and Sembenelli (2005) indeed finds that the presence of stock markets significantly stimulates VC activity.

The present chapter synthesizes our previous theoretical work and discusses the consequences of selected taxes and subsidies such as those emphasized by the EVCA benchmarking report mentioned above. They are relevant at different stages of the firm's life-cycle. The chapter explores how they impact on the quantity and quality of VC financed entrepreneurship. In particular, we examine a subsidy to start-up investment representing the various investment grants, interest subsidies and subsidies to capital expenditure in research and development which are prevalent in many countries. We explore the taxation of capital gains in new firms when sold off to new investors, the taxation of wages in occupations alternative to the pursuit of an entrepreneurial career and corporate income taxation. Our analysis indeed shows that a limited focus on the taxation of small early stage firms cuts too short. The taxation of mature firms might be as important for start-ups as the direct taxation of infant companies. The corporate income tax may well reduce entrepreneurship even though the tax is only paid by mature companies rather than young ones. The basic insight is that by reducing the value of mature firms, the corporate tax diminishes the gains from setting up new companies as well.

Our primary focus is on the consequences of taxes and subsidies for the rate of business creation and the quality of VC financing in industry equilibrium. We set up a two-period model of industry equilibrium that is rich enough to reveal the effects of taxes and subsidies on the survival probability of start-ups, IPO prices, capital investment of mature firms, and overall welfare. The core of the model is the relationship between a finance-constrained entrepreneur and a VC firm that must pay for the new firm's physical investment expenses. The firm's success rests on the entrepreneur's effort and due diligence, as is well established in the empirical literature (such as that reviewed in Rosen, 2005). It also reflects the VC's engagement and contribution to the firm as argued above and empirically documented by Gompers and Lerner (1999) and Hellmann and Puri (2000, 2002), among many others. The main functions of VC financing consist of screening, contracting and advising (see Kaplan and Stromberg, 2001, for a concise statement of the stylized facts and Kaplan and Stromberg, 2003, 2004, for detailed empirical evidence). VCs carefully screen and select business plans. They have developed sophisticated financial instruments and contractual arrangements to alleviate the problems resulting from informational

asymmetries. They add value by establishing contacts, giving strategic business advice and generally helping in the professionalization of young firms.

The chapter abstracts from screening and selection problems by assuming uniform quality among entrepreneurs. Instead, we focus on the value added role of VCs which is an important one. Hellmann and Puri (2002) show empirically that VC-backed start-ups in Silicon Valley are much faster in introducing stock option plans for high skilled personnel and in hiring a professional sales manager. Also, the presence of a VC makes it more likely that the entrepreneur is replaced by a professional CEO from outside if her lack of managerial abilities turns out to be an impediment to the firm's rapid growth. The VC's influence is particularly strong in the early phase of business development when the informational problems are the largest, but becomes insignificant later on when the firm has successfully matured. In short, the VCs add value and raise the likelihood of success by promoting the professionalization of young firms. Hellmann and Puri (2000) show that VC-backed firms introduce more radical innovations and pursue more aggressive market strategies compared with other start-ups. For example, once a VC joins the firm and provides finance, the probability of introducing the new product on the market jumps by a factor of more than three. Rapid market introduction is strategically important because the first firm on the market enjoys a first mover advantage.

Part of this superior performance of VC backed firms might result from VCs being able to locate more profitable firms than other investors, rather than adding value themselves. The issue of selection versus value added has been empirically investigated by Sorensen (2005). He indeed finds important sorting effects in the sense that the most experienced investors get the best deals. He finds that the probability of success of a firm financed by the most experienced investor in his sample is 39% which compares with a success probability of only 15% of firms financed by the least experienced investors. He then reports that sorting (the best investors getting the best deals) explains 58% of this increase in success probability while the investor's influence explains 42%. Although ignoring sorting and selection effects leads to considerable overestimation of the degree of investor influence, the value increasing role of VC financing remains a very important one.

It seems that the productive contribution of VCs to business growth is not a guaranteed matter. Apart from investor experience, it may also rest on the existence of appropriate incentives on the part of the financier. Indeed, the empirical evidence on the impact and value added of VC is less clear-cut in Europe than in the U.S. (see Bottazzi and Da Rin, 2002, for a skeptical view while Audretsch and Lehmann, 2003, paint a more positive picture). Finance theory has addressed VC incentives in terms of a double-sided moral hazard problem, where both the entrepreneur and VC must exert effort in the company



(see Holmstrom, 1982; Aghion and Tirole, 1994; Casamatta, 2003; Inderst and Mueller, 2004; Repullo and Suarez, 2004; Schmidt, 2003; and our own previous work mentioned above). Since neither party's effort is verifiable and contractible, the VC contract must be carefully crafted to provide appropriate incentives to both the entrepreneur and VC. In focusing on the real effects of VC in industry equilibrium, we postulate a particularly simple model of the entrepreneur's and VC's interaction. In this framework, a Pareto-optimal contract allocates profit shares depending on each partner's importance for the firm's success. The contract can be implemented as straight equity.

Within our simple model, the contract specifies that the VC acquires an equity stake for a price that covers at least the physical start-up costs plus possibly an upfront payment to the entrepreneur. The agreed profit sharing is chosen to optimally allocate incentives to the entrepreneur and VC in order to maximize the joint surplus to be divided among them. Although profit sharing is Pareto-optimal among the members of the team, it nevertheless implies that each party is able to appropriate only a share of the marginal gains from putting forth extra effort while she will have to bear the entire private cost of doing so. For this reason, entrepreneurial effort and VC advice tend to be too low compared to a socially efficient allocation, resulting in an overly high failure rate among start-ups.<sup>3</sup> In our set-up, no such distortion is present with respect to the rate of business creation. The literature has indeed been very skeptical toward policies that simply aim to promote the rate of business creation. In fact, it often argues for a tax rather than a subsidy to entry (cf. De Meza, 2002; see also the discussion in Cressy, 2002, and Parker, 2003). From a normative point of view, our model does not support policies to accelerate business creation either but rather argues for better quality of start-ups. A better quality reflects improved incentives for entrepreneurial effort and VC support and results in a lower failure rate among start-up firms. Our analysis supports policies that do not aim at more but rather more successful VC backed firms. There is a quality-quantity trade-off.

Most real-world policies toward young firms subsidize the cost of capital from start-up investment. Policy analysis within our model shows that these subsidies are indeed effective in stimulating entrepreneurship but are questionable from a broader welfare perspective. Since start-up subsidies are given independent of the ultimate success, they do not strengthen individual incentives for more entrepreneurial effort and VC advice. In raising the present value of a project, they make entry more attractive. Precisely because they are effective in generating entry, they tend to depress market prices and firm values which ultimately erodes the reward to private effort. Since effort is too low in private equilibrium, subsidizing start-up cost tends to reduce welfare. Capital gains taxes have an ambiguous effect on entrepreneurship while they may be quite harmful in welfare terms. Wage taxes lead individuals into entrepreneurial

careers, but likewise may be unwarranted from a welfare angle. Instead, taxes on entrepreneurs would be more sensible, leading to fewer but more successful and more valuable firms. Finally, corporate income taxes tend to be particularly harmful to the quantity and quality of VC backed entrepreneurship. Since they reduce the value of mature companies, they not only retard entry but also impair the reward to effort in start-up firms. Quite generally, any policy reducing the value of mature companies will feed back negatively on incentives within start-up firms.

To state our arguments more precisely, we now set up a stylized equilibrium model in Section 2. It will be shown how alternative policy instruments affect different behavioral margins during a firm’s life-cycle and thereby determine the quantity and quality of VC investments in industry equilibrium. Section 3 derives the impact of public policy and Section 4 concludes.

## 2. A MODEL OF START-UPS AND VENTURE CAPITAL

### 2.1. Overview

Figure 9-1 illustrates a stylized two-period model of young and mature firms. The sequence of events unfolds from left to right. At the beginning of the first period, the government defines a policy environment, consisting of the policy instruments listed at the bottom of the figure. The entrepreneurial and traditional sectors produce a perfectly substitutable output with a price normalized to unity. Production in the traditional sector is Ricardian, converting one unit of labor into  $W$  units of output, and thus paying a fixed wage  $W$ . The traditional sector absorbs all labor resources not demanded by the entrepreneurial sector.

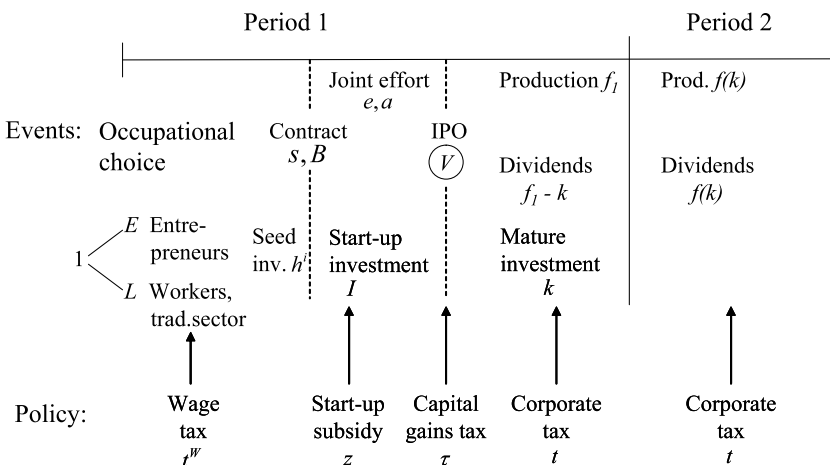


FIGURE 9-1 Events and notation.

There is a population of mass one of agents. Weighing the prospects of an entrepreneurial career against employment in the traditional sector, a mass  $E$  of agents opts for entrepreneurship to pursue their business ideas. The remaining population  $L = 1 - E$  chooses employment. The occupational choice decision of individuals thus shifts production to one or the other sector. In the second period, output is supplied by entrepreneurial firms only, traditional firms being inactive.

An entrepreneur must first undertake a “seed investment” to turn her idea into a project and develop a business plan which can be presented to a financier. For this purpose, individual  $i$  needs to incur a (nonpecuniary) cost  $h^i$ , capturing the time and effort required in the process. Individuals are assumed to differ in their basic inventiveness. Some create their project at low cost, while others have to put in more effort. Lacking own resources to start the firm, an entrepreneur proposes a deal to a VC firm to finance and advise the venture. When accepting the contract, the VC acquires a profit share  $1 - s$ , leaving a share to the entrepreneur, against a total price  $B + (1 - z)I$  that covers at least the private start-up investment  $I$  net of a possible government subsidy  $z$  plus an extra upfront payment  $B$  to the entrepreneur. The parameters  $s$  and  $B$  of the contract are optimally chosen to reflect the relative importance of the entrepreneur’s and the VC’s contribution to the firm’s success during the start-up phase. The contract must also be sufficiently attractive to induce the entrepreneur to give up the alternative career.

Having specified the terms of the contract, the firm is started up with a fixed capital investment  $I$ . The venture is risky. Both the entrepreneur and the VC must put in effort to enhance the firm’s chances. The likelihood of success is specified as  $p = p(e, a)$  and depends on entrepreneurial effort  $e$  and VC advice  $a$ . If a venture succeeds, production starts, and the firm can be sold to new investors, possibly at an IPO, for a price  $V$ . At that point in time, the capital gains tax at rate  $\tau$  is due and cuts into the privately earned gains during the start-up period.<sup>4</sup> If it fails (with probability  $1 - p$ ), the firm will be shut down without any production and revenues whatsoever. When firms successfully mature to production stage, they produce  $f_1$  for the remainder of the first period. A part  $k$  of this production is retained and invested internally to accumulate capital, while the residual is distributed as dividends to owners. In the second period, production  $f(k)$  is continued at a level depending on mature firm investment  $k$ . Revenues net of the corporate tax at rate  $t$  are paid out to owners. The capital stock  $k$  is assumed to depreciate in full over the second period. Depending on the level of wages or entrepreneurial income received, and on the market rate of interest, individuals choose optimal life-cycle consumption.

The policy instruments to be investigated are:  $t^W$  a tax on wage income;  $t$  a corporate income tax on mature firms;  $\tau$  a capital gains tax on new firms, levied symmetrically on entrepreneurs and VCs; and  $z$  a subsidy to start-up

investment. A fraction  $\theta$ ,  $0 \leq \theta \leq 1$ , of mature firm investment  $k$  can be expensed in the first period from the corporate income tax, even though capital only depreciates in the second period. The remaining part  $1 - \theta$  is deducted from the tax base in the second period. Government budget imbalances are offset with lump-sum taxes or transfers in the same period.

The model is solved by backward induction, in the reverse order of Figure 9-1. Most of the elements of the model are as in Keuschnigg and Nielsen (2004b); however, we do amend the framework in that article to enable an analysis of corporate income taxation. The full details of our model and its solution are available in the working paper version Keuschnigg and Nielsen (2004c). Below, we shall limit our remarks to those model elements which are most crucial to understanding the effects of taxes/subsidies on entrepreneurship and welfare.

## 2.2. Mature Firm Value and Investment

Consider first the investment decision and the valuation of mature firms. A mature firm is assumed to pay net of tax dividends  $\chi_1 = (1 - t)f_1 - (1 - \theta t)k$  and  $\chi_2 = (1 - t)f(k) + (1 - \theta)tk$ , where  $f_1$  is a fixed amount of first period output and  $f(k)$  is second period output,  $f(\cdot)$  being a standard concave production function. A part  $\theta$  of mature firm investment is immediately expensed against the corporation tax; the remaining part reduces the tax bill next period. This definition of dividends assumes internal investment finance. At IPO, the value  $V$  of a mature firm reflects the present value of the net dividend flow,  $V = \chi_1 + \chi_2/R$ . The gross interest factor is  $R = 1 + r$ ,  $u$  is the user cost of capital and  $V_1$  denotes that part of firm value which is optimized with respect to mature firm investment:

$$V = (1 - t)f_1 + V_1, \quad V_1 = \frac{(1 - t)[f(k) - uk]}{R}, \quad u \equiv \frac{(1 - \theta t)R - (1 - \theta)t}{1 - t}. \quad (1)$$

Maximizing firm value with respect to investment  $k$  yields:

$$f'(k) = u \Rightarrow \frac{dk}{dR} < 0, \quad \frac{dk}{dt} \leq 0, \quad \frac{dk}{d\theta} > 0. \quad (2)$$

Since the corporate income tax raises the cost of capital, it reduces mature firm investment. If  $\theta = 1$ , however, so that capital investment can be immediately expensed, the corporate income tax becomes a cash flow tax, neutral to investment (presuming a positive corporate income tax). An increase in the rate of immediate investment expensing promotes investment if the tax rate is positive. Finally, a rise in the interest rate tends to lower investment in mature firms.

Using the envelope theorem, the effects of taxes on mature firm values are:

$$\begin{aligned} \frac{dV}{dR} &= -\frac{V_1 + (1 - \theta t)k}{R}, & \frac{dV}{d\theta} &= t\frac{rk}{R}, \\ \frac{dV}{dt} &= -\frac{V}{1-t} \left[ 1 + \frac{(1 - \theta)rk}{RV} \right]. \end{aligned} \tag{3}$$

The IPO value will be negatively affected by increases in both the corporate income tax and the interest rate, while a rise in the expensing parameter stimulates firm value provided that the corporate tax is positive. It is important to note that the corporate income tax always reduces mature firm value irrespective of how it affects their investment decisions. The value of a mature firm is the ultimate reward for starting up a firm in the first place. Thereby, mature firm value determines the incentives of entrepreneurs and their financiers in the preceding start-up phase.

### 2.3. VC Financed Start-ups

The entrepreneur’s interest in starting a firm and the VC’s willingness to finance such a venture depend on the value of the firm at the beginning of the start-up phase and the terms of the financial contract. Since young firms survive the start-up phase only with probability  $p$ , the expected value of going ahead with the project is  $pV$  minus various costs. A start-up succeeds with probability  $p$ , leaving a value of  $V$ , and fails with  $1 - p$ , leaving nothing. The success probability  $p = p(e, a)$  increases with the entrepreneur’s effort  $e$  and the VC’s managerial support  $a$ . Both inputs are subject to decreasing returns. For simplicity, we specify:

$$p = p(e, a) = e^\varepsilon a^\alpha, \quad \varepsilon + \alpha < 1. \tag{4}$$

Vcs and entrepreneurs share expected firm value by means of a simple equity contract where the VC acquires a profit share  $1 - s$  at a total price of  $B + (1 - z)I$ . The subsidy  $z$  stands for the variety of government programs such as interest subsidies, credit guarantees or outright investment subsidies that are all intended to reduce the cost of capital to young entrepreneurial firms. While the part  $(1 - z)I$  of the VC’s payment is spent on capital equipment, the entrepreneur possibly retains an upfront compensation  $B$  that makes her more willing to forego alternative opportunities. With such a contract that will be optimally chosen, the expected incomes accruing to entrepreneurs, VCs and the

government are:

$$\begin{aligned}
 \pi^E &= (1 - \tau)[spV + B], \\
 \pi^F &= (1 - \tau)[(1 - s)pV - B - (1 - z)I], \\
 \pi^G &= \tau[pV - (1 - z)I] - zI, \\
 \pi &= \pi^E + \pi^F + \pi^G = pV - I.
 \end{aligned} \tag{5}$$

Note that  $\tau$  stands for the uniform capital gains tax on VCs and entrepreneurs. The government's surplus  $\pi^G$  corresponds to the *net* tax revenue extracted from the project.

To raise the chances of the firm's success, the entrepreneur and the VC must incur effort costs  $\beta e$  and  $\gamma a$ , respectively. Since these costs are non-pecuniary, effort cannot be verified and, therefore, cannot be secured by explicit contracting. The contract must thus enable both the entrepreneur and the VC to participate in the upside potential of the firm by allowing them to share in the extra profit derived from their effort. Define the entrepreneur's profit net of effort cost as  $\Omega^E \equiv \pi^E - \beta e$  and the VC's surplus per venture as  $\Omega^F \equiv \pi^F - \gamma a$ . The entrepreneur proposes a contract that maximizes her surplus  $\Omega^E$  subject to two considerations. First, the contract must be sufficiently profitable for the VC to assure her willingness to finance the venture ( $\Omega^F \geq 0$ , participation constraint  $PC^F$ ). Second, the entrepreneur must anticipate how the proposed profit sharing rule will affect her own and the VC's incentives to provide effort *after* the contract is signed (incentive compatibility constraints  $IC^E$  and  $IC^F$ ). Formally, the entrepreneur's problem is:

$$\Omega^E = \max_{s, B} (1 - \tau)[p(e, a)sV + B] - \beta e \quad s.t. \tag{6}$$

$$PC^F: \Omega^F = (1 - \tau)[p(e, a)(1 - s)V - B - (1 - z)I] - \gamma a \geq 0, \tag{i}$$

$$IC^E: \Omega_e^E = p_e(e, a)(1 - \tau)sV - \beta = 0, \tag{ii}$$

$$IC^F: \Omega_a^F = p_a(e, a)(1 - \tau)(1 - s)V - \gamma = 0. \tag{iii}$$

At effort stage, where the agreed profit share  $s$  is already fixed, optimal levels of efforts are determined by the two incentive compatibility constraints. Figure 9-2 illustrates the simultaneous choice of effort, using the functional form for  $p(e, a)$  in (4). Both reaction curves  $e(a)$  and  $a(e)$  are positively sloped, implying that entrepreneurial effort and VC advice are strategic complements. According to Figure 9-2, a larger expected IPO value boosts both the entrepreneur's effort and the VC's managerial support and thereby raises the firm's survival chances. An increase in the symmetric capital gains tax reduces the

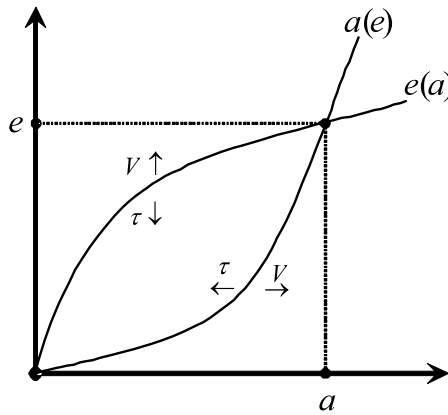


FIGURE 9-2 Effort and advice.

reward for effort and yields the opposite effects. Finally, inspection of (6.ii–iii) shows how the contract allocates incentives for effort by specifying the profit sharing rule. The higher is the entrepreneur’s share  $s$ , the larger are her marginal returns to effort, and the more effort she will supply. At the same time, a higher  $s$  is at the expense of the VC’s share  $1 - s$ . In this case, the VC will be less keen to support the venture with her advice and contacts which reduces the firm’s survival chances.

Anticipating effort choices during the start-up phase, the entrepreneur proposes a deal that maximizes her own surplus but at the same time assures that the VC is willing to finance the investment expenditure. In assuming competitive VCs, we allocate all bargaining power to the entrepreneur. Accordingly, the VC’s surplus per venture,  $\Omega^F \equiv \pi^F - \gamma a$ , is squeezed to zero.<sup>5</sup> The entrepreneur can raise her own expected profit by keeping either a larger share  $s$  or demanding a higher upfront payment  $B$  by asking for a higher price for the VC’s profit share. Note a fundamental difference between the two instruments  $s$  and  $B$ . Claiming a higher  $s$  reduces the VC’s share and destroys her incentives to add value, while the upfront payment  $B$  does not. The latter merely redistributes lump-sum across the two parties. The entrepreneur will therefore first choose to maximize joint surplus. Having found this Pareto-optimal share  $s$ , she then requests a maximum upfront payment  $B$  that allows the VC no more than to break even. In this way, the entrepreneur acquires the entire joint surplus  $\Omega = \Omega^E + \Omega^F$ . Substituting  $B$  from (6.i) into (6) yields the entrepreneur’s problem of choosing  $s$ , anticipating the incentive effects for later stage effort and as determined by (6.ii–iii):

$$\Omega \left[ \underset{+}{V}, \underset{-}{\tau}, \underset{+}{z} \right] = \max_s (1 - \tau) [p(e, a)V - (1 - z)I] - \gamma a - \beta e \quad s.t. \text{ (6.ii–iii)}. \quad (7)$$

With a symmetric capital gains tax on both entrepreneurs and VCs, the Pareto-optimal profit share  $s$  becomes independent of taxes and of venture returns  $V$ . We can thus take  $s$  as a fixed constant, beyond the influence of policy.<sup>6</sup>

#### 2.4. Entry to Entrepreneurship

An entrepreneur's expected surplus is the utility difference between entrepreneurship and employment and reflects not only income differences but also various effort costs. First, seed investment is interpreted as a nonpecuniary private research effort which is required to prepare a business plan. Agents are taken to be distributed uniformly in the unit interval with respect to research ability and associated effort cost,  $h^i$  in monetary equivalent. Once this effort is sunk and the alternative wage income is foregone, all start-up firms are assumed to be of uniform quality and to yield the same remaining surplus  $\Omega$ .<sup>7</sup> The expected surplus must be sufficiently large to compensate entrepreneurs for any foregone outside opportunity  $(1 - t^W)W$ , and the initial effort cost  $h^i = h \cdot i$  during the seed phase prior to VC finance. Entry of entrepreneurs occurs as long as  $\Omega - h \cdot i > (1 - t^W)W$ , until the marginal entrepreneur just breaks even.<sup>8</sup> Entrepreneurial entry is, thus, governed by:

$$\Omega[V, \tau, z] = hE + (1 - t^W)W. \quad (8)$$

On the left-hand side, entrepreneurial surplus increases in mature firm value  $V$  and start-up subsidy  $z$ , but falls in the tax rate  $\tau$  that diminishes the expected capital gains of start-up investment. Note that start-up firms prior to production stage do not generate any profits and therefore do not pay corporate income tax. However, since corporate taxes reduce mature firm value as in (3), they nevertheless reduce the expected reward from start-up entrepreneurship.

Figure 9-3 illustrates the relation between venture returns and the number of entrepreneurs. A higher venture return  $V$ , consisting of a higher IPO value of a maturing firm, raises the returns to start-up activity and leads more agents to choose an entrepreneurial career. The other policy effects are directly inferred from the Figure 9-3.

#### 2.5. Industry Equilibrium

Equilibrium in the venture capital based industry reflects the demand for and supply of entrepreneurial firms.<sup>9</sup> While supply of start-up entrepreneurs result from the agent's occupational choice as illustrated in Figure 9-3, demand stems from the economy's need of mature firm output. We have chosen a very simple formulation of demand. Depending on their income, workers and



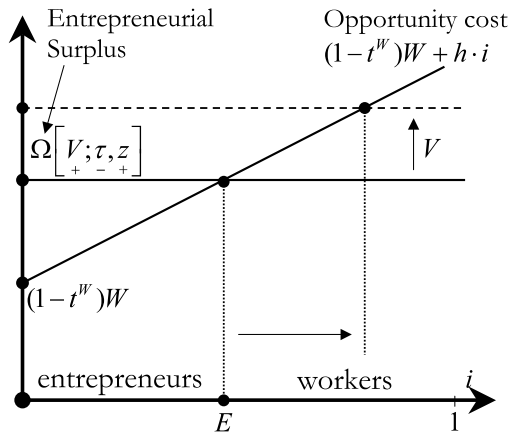


FIGURE 9-3 The start-up decision.

entrepreneurs spend on first and save for second period consumption. In postulating linear and separable preferences, we cut out any complicating income effects. Thus, demand for second period consumption exclusively depends on the interest rate. With a higher interest rate, agents save more and demand more of second period consumption  $D$ , with  $D'(R) > 0$ . Therefore, demand for mature firm output, equal to  $f(k)$  per firm as indicated in Figure 9-1, increases with the interest rate. As this must be covered by successful start-up firms, it ultimately creates demand for more entrepreneurs. By the law of large numbers with independent risks, the number of mature firms is  $N = pE$ . Demand for entrepreneurship thus reflects demand for second-period consumption  $D$  and the number and size of mature firms. Hence:

$$D = f(k) \cdot p(e, a) \cdot E^D. \tag{9}$$

On the right-hand side, second-period production is the product of output per firm, the number of start-ups (entrepreneurs) and the success rate of start-ups. The equation can be solved for the number of entrepreneurs demanded  $E^D$  which is easily shown to be an increasing function of the interest rate as indicated in Figure 9-4. First, a higher interest rate raises demand for entrepreneurs because it boosts savings and thus demand for mature firm output  $D$ . Second, a higher interest rate depresses mature firm investment  $k$  and thereby reduces output  $f(k)$  per mature firm. With smaller size, more of them are required to serve the market which again contributes to more demand for entrepreneurship. Third, a higher interest rate depresses the value of successful new firms  $V$  as in (3). Lower venture returns dampen the incentives for joint effort as indicated in Figure 9-2 and thereby reduce the success rate  $p(e, a)$  of new firms so that more of them must start up to serve the market. Note

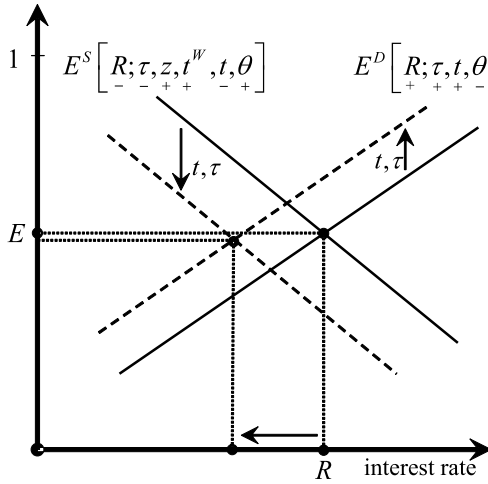


FIGURE 9-4 *Equilibrium entrepreneurship.*

that both  $e$  and  $a$  and thus  $p$  symmetrically increase with  $V$  but fall with  $\tau$  as indicated in Figure 9-2. Consider further the factors determining mature firm value as listed in (3). Put all the information together and solve (9) for the required number of entrepreneurs. It is then seen how demand for start-up entrepreneurship increases with the interest rate and shifts with tax rates as indicated in Figure 9-4. Formally, the demand for entrepreneurs is

$$E^D \left[ R; t, \theta, \tau \right] = \frac{1}{p(V(R, t, \theta), \tau)} \cdot \frac{D(R)}{f(k(R, t, \theta))}. \tag{10}$$

Tax parameters shift the demand schedule for entrepreneurs. An increase in the corporate income tax  $t$  lowers investment in all firms, necessitating more firms to meet demand. Moreover, the higher tax lowers the price of successful new firms. This depresses efforts and the probability of success of new firms so that more of them have to start up to meet demand for second period goods. As a result, a higher corporate income tax moves the demand schedule up. The opposite effects can be registered for an increase in the depreciation parameter  $\theta$ . Finally, a higher capital gains tax  $\tau$  raises the demand for entrepreneurship since the tax reduces the returns to effort and thus cuts into the success rate, so that more firms must be created to satisfy any given demand for second period output. Again, the demand schedule shifts up.

The supply schedule in Figure 9-4 results from the occupational choice decision in (8). It slopes down with the interest rate. Since an increase in the interest rate lowers venture returns  $V$ , the entrepreneur's surplus  $\Omega$  is reduced, so that fewer individuals find it worthwhile to incur the seed investment  $h^i = h \cdot i$

as is illustrated in Figure 9-3. More formally, the free entry condition (8) yields:

$$E^S \left[ R; \tau, z, t^W, t, \theta \right], \quad \frac{dE^S}{dR} = \frac{1}{h} \cdot \frac{\partial \Omega}{\partial V} \frac{dV}{dR} < 0. \quad (11)$$

Apart from the negative interest rate effect on the supply of entrepreneurship, the capital gains tax likewise tends to reduce entrepreneurship on account of its negative effect on entrepreneurial surplus. In contrast, a higher start-up capital subsidy and a higher wage tax both stimulate entrepreneurship, since they respectively boost the surplus created by entrepreneurial firms and lower the opportunity cost of entering into entrepreneurship. Finally, the corporation tax subtracts from mature firm value  $V$  and thereby the reward to entrepreneurship. Tax depreciation adds value and consequently encourages start-up activity which shifts up the entrepreneurial supply schedule.

Equating demand and supply yields the equilibrium level of start-up entrepreneurship with the gross interest rate  $R$  being the equilibrating price. Figure 9-4 illustrates this. The comparative static effects of tax/subsidy changes follow from simple graphical arguments and will be discussed in the following section. The model also helps to develop a welfare based policy perspective toward the VC industry. Our welfare measure is the sum of individual utilities. It fully takes account of all nonpecuniary effort costs and also reflects the government budget constraint. It thereby appropriately considers the cost of public funds that might be channeled to the VC industry.

## 2.6. Efficiency

Our previous theoretical work has identified the double moral hazard in the relationship between the entrepreneur and the VC as a source of inefficiency while all other behavioral margins such as occupational choice and entrepreneurial entry are free of distortions.<sup>10</sup> Since efforts are assumed not verifiable and not contractible, neither the entrepreneur nor the VC is able to commit to first best effort but will choose their inputs according to the incentive constraints (6.ii–iii). Both agents must share the return on their effort within the team, but must fully bear their own cost, implying that entrepreneurial effort and VC advice are too low in the private equilibrium.<sup>11</sup> As a consequence, and starting from an untaxed state, welfare is positively related to marginal increases in entrepreneurial effort or VC advice. Even small taxes can thus give rise to first order welfare changes. To obtain strictly positive welfare gains relative to the *laissez faire* equilibrium, one must look for policies that boost entrepreneurial effort and VC support rather than entry. In our framework, policy should not so much focus on the quantity but rather on the quality of VC financing.

How robust is this bias toward low entrepreneurial effort and managerial support which contributes to lower quality of VC financing? Our assumption that the VC and entrepreneur jointly determine the success probability and must exercise effort simultaneously, is important. Schmidt (2003), for example, assumes sequential efforts where in a first phase only the entrepreneur's effort is required while in a second stage the further increase in the value of the firm depends exclusively on the VC's managerial input. With this sequential effort choice, Schmidt (2003) is able to explain the use of convertible debt. In particular, he shows that convertible securities serve to obtain a first best outcome in his framework. While convertible debt is certainly a more flexible financial instrument and may allow parties to attain a superior outcome relative to straight equity finance, the first best result hinges critically on the fact that efforts are never required simultaneously but only sequentially, see Schmidt (2003, section III.G). Different from Schmidt, we stress the fact that the entrepreneur's effort is critical throughout the company's life. In reality, most business failures are ultimately due to some entrepreneurial management mistake. When the joint efforts of entrepreneur and financier overlap and are required simultaneously, the possible advantage of convertible securities relative to (mixed) equity contracts is reduced. In this case, the basic inefficiency noted above emerges again.

Compared to our simplified model, real world VC contracts contain of course many additional elements such as staging and syndication of venture capital investments (Lerner, 1994; Cornelli and Yosha, 2003) or the use of control rights (Hellmann, 1998; Gompers and Lerner, 1996). This extra contractual flexibility should make VC contracting more efficient in reality. However, these non-monetary incentives may be considered more like complements rather than substitutes to the incentives provided in a financial contract (see Hart, 2001). Neglecting them in our analysis should thus not affect the basic policy conclusions.

### 3. POLICY AND THE VENTURE CAPITAL SECTOR

The model presented in the preceding section is well suited to study how fiscal policy might affect the joint efforts of entrepreneurs and VCs in new firms, the success probability of these, the level of entrepreneurship, venture returns, and welfare. The following analysis will emphasize intuitive explanations. For a more formal analysis of the proposed policy experiments we refer to Keuschnigg and Nielsen (2004b, c). Table 9-1 provides an overview of the main results.

TABLE 9-1 *Effects of taxes and subsidies*

| Type of tax                            | $R$                 | $E$   | $N$ | $V$ | $e$ | $a$ | $U^*$ |
|--|---------------------|-------|-----|-----|-----|-----|-------|
|  | Mature firms        |       |     |     |     |     |       |
| Corporate tax <sup>#</sup> $t$         | -                   | -     | -   | -   | -   | -   | -     |
| Tax depreciation <sup>#</sup> $\theta$ | +                   | +     | +   | +   | +   | +   | +     |
|  | Young firms         |       |     |     |     |     |       |
| Capital gains tax $\tau$               | -                   | $\pm$ | -   | -   | -   | -   | -     |
| Start-up subsidy $z$                   | +                   | +     | +   | -   | -   | -   | -     |
|  | Occupational choice |       |     |     |     |     |       |
| Wage tax $t^W$                         | +                   | +     | +   | -   | -   | -   | -     |

Note:  $R$  interest factor,  $E$  young firms,  $N = p(e, a)E$  mature firms,  $V$  value of mature firm,  $e$  entrepreneurial effort,  $a$  venture capital advice,  $U^*$  welfare.

<sup>#</sup> The change in the interest rate is unambiguous.

### 3.1. Corporate Taxation

The effects of taxes are best understood in terms of the demand and supply curves for entrepreneurial firms. The supply side reflects the occupational choice of entrepreneurs. An increase in the corporate tax directly reduces the value of a mature firm which diminishes the entrepreneurs' surplus from creating a new one. Fewer entrepreneurs will want to incur the opportunity costs and give up alternative wages. Accordingly, the supply curve in Figure 9-4 shifts down.

For any given size of the output market, the demand for entrepreneurship follows from the number of mature firms  $N$  needed to supply the market,  $D = f(k)N$ . A first policy effect derives from its impact on output per firm. Since the corporate tax impairs expansion investment and thereby erodes output per mature firm, a larger number  $N$  of them is needed to serve the market which enhances demand for entrepreneurship. Second, since only a fraction  $p$  of new companies actually mature to production stage,  $N = pE$ , the number of young firms must necessarily be larger than the mass of established businesses. As the corporate tax diminishes the value of a mature firm, it erodes the incentives for entrepreneurial effort and managerial advice and leads to an increased rate of business failure. Everything else equal, more new firms must be started for any given mass of mature firms serving the demand for second period output. Both effects shift up the demand schedule in Figure 9-4.

To eliminate the resulting excess demand for entrepreneurship, the interest rate must fall. Along the supply curve, entrepreneurship picks up, since a lower rate of interest raises mature firm value which creates a larger surplus from business creation and thereby attracts more entrepreneurs to set up their own firm. Turning to the demand side, we find that a lower interest rate depresses savings and demand for second period output. Moreover, a lower interest rate boosts expansion investment, making mature firms bigger and

thereby requiring fewer of them to serve the market. Further, the increased firm value boosts joint effort and thereby survival rates so that fewer start-ups are needed for any given number of mature firms. All three effects (i.e., smaller market, bigger mature firms and a higher survival rate of young firms) add up to reduce demand for entrepreneurship along the demand curve.

Apparently, the equilibrium effect on entrepreneurship is ambiguous when both curves shift as illustrated in Figure 9-4. In Keuschnigg and Nielsen (2004c) we derive a sufficient condition for the net effect to be negative as stated in Table 9-1.<sup>12</sup> The corporate income tax discourages entrepreneurial effort and VC support and thereby contributes to a higher rate of business failure. While a falling interest rate boosts firm value, a higher tax reduces it. The direct tax effect dominates to reduce the value of a mature firm and thereby diminishes the returns to effort during the start-up phase. The corporate tax thereby leads to a first order welfare loss since efforts are already too low and the rate of business failure too high in the market equilibrium. This first order welfare loss is much more severe than the tax distortion of mature firm investment which results only in a second order welfare loss that would disappear for small taxes. We summarize:

**PROPOSITION 1 (Corporate Tax on Mature Firms).** *(a) The corporate income tax reduces market size and the equilibrium interest rate. (b) The corporate tax decreases the number of start-up and mature firms and lowers firm value. (c) It impairs incentives for effort and advice and reduces the success probability. (d) A small tax increase entails a first order welfare loss.*

The tax allowance  $\theta$ , that is, the share of investment outlays immediately deductible from the current tax base, allows us to portray different systems of corporate income taxation. Note that we have assumed full depreciation of capital in each production period. Setting the tax allowance to zero corresponds to a Schanz-Haig-Simons corporate income tax with tax depreciation equal to economic depreciation in the second period; see Section 2.2. In contrast, immediate expensing of investment outlays corresponding to  $\theta = 1$  represents a cash flow tax. Having undertaken an immediate write-off prevents, of course, further tax depreciation in the second period when capital actually depreciates economically. The cash flow tax is well known to be neutral with respect to investment, resulting in a *marginal effective tax rate* on expansion investment equal to zero. In this case, the user cost of capital in (1) exclusively depends on the rate of interest but is independent of the tax rate. However, the *average effective tax rate* of the cash flow tax (i.e., the share of corporate income paid in tax) is strictly greater than zero since it continues to tax economic rents unrelated to the returns on marginal investments. Notwithstanding the neutrality of the cash flow tax with respect to marginal expansion investment, the tax

burden is capitalized in a lower firm value. In reducing the IPO price, the cash flow tax does distort against discrete start-up investment. It also impairs the incentives of entrepreneurs and VCs to engage in their firms and thereby contributes to more frequent business failure. Given that joint efforts are already too low from a social perspective, the cash flow tax diminishes welfare and efficiency as in Proposition 1.

Consider now the effects of more favorable tax depreciation or tax allowance for expansion investment, that is, an increase in  $\theta$ , and keep the tax rate constant. Of course, investment expensing is valuable only if the tax rate is positive already. A more generous allowance promotes expansion investment and, by reducing the average effective tax rate, boosts firm value. Given a larger value to be realized at IPO, entrepreneurs can expect a larger surplus from business creation and will accordingly start businesses more often. In consequence, the supply schedule in Figure 9-4 for young entrepreneurial firms shifts up (to be drawn by the reader). At the same time, the expectation of larger IPO values invigorates the joint effort in the start-up phase and contributes to improved survival rates. With higher survival chances, fewer firms need to be started if any given number of them must reach the production stage. The increased tax allowance further raises expansion investment and production in mature firms which likewise reduces the demand for entrepreneurship. The demand schedule in Figure 9-4 thus moves down for both reasons.

Obviously, to eliminate the resulting net supply of entrepreneurial firms, the interest rate must rise to force exit. Although a higher interest rate erodes firm values, it does not overturn the positive direct effect of the tax allowance. The net effect is an increase in IPO value which boosts the return to effort and encourages VCs to advise more intensively. Start-up firms accordingly benefit from this extra effort in terms of improved survival chances. Given that joint effort is too low initially, the tax allowance results in a first order welfare gain. Finally, the rate of business creation and the number of mature firms result from offsetting influences. First, the higher equilibrium interest rate reflects larger market size due to higher demand for second period output which expands the demand for both types of firms. Second, the tax allowance boosts marginal investments and makes mature firms bigger. The market supports a smaller number of them which negatively feeds back on the rate of business creation as well. The analysis in Keuschnigg and Nielsen (2004c) finds the net effect to be positive. Third, given that start-ups are more likely to mature to production stage, fewer of them are needed for any given number of firms on the product market. Again, our computations report a net positive effect.

**PROPOSITION 2 (Tax Allowance for Expansion Investment).** *(a) With a corporate tax in place, a more generous tax allowance for expansion investment raises equilibrium interest and boosts market size. (b) The tax allowance boosts*

*firm values and raises the number of young and mature firms. (c) In raising firm values, the allowance sharpens incentives for effort and advice and boosts the success probability. (d) By raising mature firm values, the tax allowance stimulates effort and leads to first order welfare gains.*

### 3.2. Capital Gains Taxes and New Firms

The immediate effect of a capital gains tax on young firms, given expected IPO values  $V$ , is to subtract from returns to effort and advice. The tax does not directly affect mature firm value which is exclusively determined by corporate taxes and the market interest rate. As illustrated in Figure 9-2, the tax discourages entrepreneurial effort and managerial advice and consequently results in a higher failure rate among start-up firms. The increased risk affects both the supply and demand schedules for entrepreneurship. In reducing the expected surplus from entrepreneurship, fewer agents find it worthwhile to start their own firm. The supply curve thus shifts down as indicated in Figure 9-4. On the demand side, the tax has no direct impact on market size and expansion investment of mature firms. However, on account of the reduced success probability of young firms, more entrepreneurs are required for any given number of firms to mature to production stage. The demand curve thus shifts up.

In face of the emerging excess demand for entrepreneurship, the interest rate must fall to re-establish equilibrium. The lower interest rate leads to lower savings and second period demand for goods, shrinking the market size. It also encourages mature firm investment and boosts firm values which, in turn, stimulate the returns to joint effort in the start-up phase. For all three reasons — smaller market size, larger mature firms, and a higher survival rate among start-ups—the demand for new firms falls and reduces entrepreneurship. The increase in firm values on the other hand boosts entrepreneurial surplus and stimulates the supply of new entrepreneurs along the supply curve. The net effect on the equilibrium number of start-up entrepreneurs remains ambiguous. The ambiguity arises despite the tax leading to a smaller number of mature firms. Mature firms also grow bigger since the falling interest rate spurs expansion investment. More entrepreneurs might nevertheless be needed since a lower success rate requires more start-ups for enough of them to mature to production stage.

While the tax discourages joint effort for any given IPO value  $V$ , the falling interest rate raises mature firm value and thereby sharpens incentives for effort. In Keuschnigg and Nielsen (2004c) we show that this price adjustment cannot dominate over the direct tax effect, implying lower effort and VC support and, hence, a lower success rate in equilibrium. The reduction in entrepreneurial effort and VC support leads to a welfare loss.



PROPOSITION 3 (Capital Gains Tax on Start-up Firms). (a) *A symmetric capital gains tax reduces the interest rate and market size. (b) On account of a lower rate of interest, the tax boosts mature firm value, raises expansion investment but reduces the number of mature firms. The change in the number of start-up firms is ambiguous. (c) The tax impairs incentives for effort and advice and reduces the survival probability. (d) Introducing a small capital gains tax on start-up firms entails a first order welfare loss.*

A corollary of this proposition is that a small negative capital gains tax—or a revenue subsidy—for young firms will encourage effort and VC support and thereby contribute to higher welfare. However, a possible tax break in capital gains taxation must be limited to young VC-backed firms only. We have also assumed full loss offset in capital gains taxation. The results on the capital gains tax are robust to restrictions on loss offset. Interestingly, the loss offset limitation can itself strengthen incentives for VC support in that the tax penalty arising from a limited loss offset makes business failure more costly (Keuschnigg and Nielsen, 2003b).

### 3.3. A Subsidy to the Cost of Capital

Most real world policies to encourage business creation allow for interest subsidies, loan guarantees to facilitate access to cheaper bank loans, or direct subsidies to investment spending. All these measures subsidize the cost of capital and are largely unrelated to firm performance. They can thus be understood as a subsidy to the cost of start-up investment, captured by  $z$  in our model. The only direct effect of an increase in the investment subsidy is to raise the entrepreneur's surplus from starting the firm and thereby to encourage entry; see (7) and (11). There are no other direct effects neither on effort and advice nor on the demand for start-up firms. In Figure 9-4, the subsidy thus shifts up the supply schedule, creating excess supply of entrepreneurs. The adjustment mechanism is well known by now. The interest rate must rise to stimulate savings and demand for second period output which leads to more demand for mature and young firms. At the same time, the increase in the interest rate erodes firm value and entrepreneurial surplus which cuts back on entry and supply of new firms. The new equilibrium is characterized by a higher interest rate, larger market size and supports a larger number of entrepreneurs and mature firms. The higher interest rate retards mature firm investment and erodes firm values; see Table 9-1.

The undesirable side effect of start-up subsidies is that they impair incentives for entrepreneurial effort and VC advice. The success probability correspondingly declines. The more successful these subsidies are in stimulating entry, the more likely should be the decline in venture returns and the

stronger the negative welfare consequences. Note, however, that the welfare loss results from a general equilibrium effect rather than any direct impact. In a small open economy with a fixed interest rate, mature firm value should remain constant. In this case, the incentives for joint effort would remain untarnished and the subsidy would only produce increased entry. Since the entry margin is not distorted, the subsidy would entail a zero welfare effect in this case.<sup>13</sup>

**PROPOSITION 4 (Capital Subsidy to Start-ups).** *(a) A subsidy to start-up capital cost raises the interest rate and expands market size. (b) The subsidy expands the number of young and mature firms, but erodes mature firm value. (c) It impairs incentives for effort and VC advice and reduces the survival rate. (d) Introducing a small subsidy entails a first order welfare loss.*

The fact that a start-up subsidy and the capital gains tax both reduce welfare suggests the following strategy that would contribute to a more active VC industry, yet avoid any high cost to the general tax payer. Impose a tax  $z < 0$  on start-up investment cost and use the proceeds to finance a narrow tax break  $\tau < 0$  on capital gains to young VC backed firms. Since the entrepreneur is wealth constrained, the start-up tax must be paid by the VC who should have no difficulty in raising capital and who will share the revenue subsidy with the entrepreneur when the venture succeeds. Being self-financed, the policy provides a net tax or subsidy equal to zero. A small start-up tax thus finances a cut in the capital gains tax rate by  $(pV - I)d\tau = Idz$ .

Consider first the direct impact for a given mature firm value  $V$ .<sup>14</sup> The direct effects on entrepreneurial surplus from the investment tax and from the revenue subsidy exactly cancel out because the policy is constructed to be self-financing. However, the tax break on  $\tau$  strengthens incentives, thereby boosting joint effort as illustrated in Figure 9-2, and consequently increases the success rate as well. As a result, the project surplus increases and encourages entry of entrepreneurial firms. The supply schedule in Figure 9-4 shifts up. At the same time and for any given  $V$ , the tax cut  $\tau$  reduces the demand for entrepreneurship because it makes start-ups more successful by inducing more effort. Fewer firms are needed to satisfy goods demand if more of them mature to production stage. The demand schedule shifts down. The equilibrium effect on entrepreneurship remains ambiguous, but the interest rate goes up to close the gap between demand and supply. Furthermore, it is easily shown that net venture values  $(1 - \tau)V$  increase on account of the tax cut. Accordingly, the self-financing policy stimulates joint effort and raises the survival rate in equilibrium as well. This brings about an improvement in welfare.<sup>15</sup>

Our framework hence essentially implies that public policy should not aim at more, but at more successful VC backed firms. Policy should not aim at the volume, but at the quality of VC investments. This conforms quite well with

the conclusions of Bottazzi and Da Rin (2002) and Hege et al. (2003) about VC in Europe. They argue that in Europe VC has expanded quite impressively over the last decade, but the impact on firm performance seemingly remained rather limited. If anything, this calls for a policy that sharpens incentives for more entrepreneurial effort and more active VC involvement. In our framework, the entry margin is undistorted, but the double moral hazard between entrepreneurs and VCs works to erode incentives for value creating effort. While in many countries current policy *vis-à-vis* start-up firms essentially consists in a series of subsidies to investment in these firms, coupled with taxation of capital gains, our analysis suggests that scaling down these subsidies and using the budget savings to finance a narrow tax break for capital gains on VC backed investments would be beneficial.

### 3.4. Wage Taxation

The rate of business creation depends not only on the surplus created by new entrepreneurial firms but also on the entrepreneurs' alternative career prospects. For this reason, wage taxation is quite relevant for start-up activity as the empirical literature mentioned in the Introduction emphasizes. The implications of wage taxation in our model are easily inferred. The wage tax exclusively influences the occupational choice decision. In reducing the opportunity cost of entrepreneurship, it stimulates entry of new entrepreneurs and thereby shifts up the supply schedule in Figure 9-4. To equilibrate demand and supply of new entrepreneurial firms, the interest rate must rise. The higher interest rate presses down the value of new firms at IPO. Lower venture returns, in turn, hurt effort and advice in start-up firms, harm their survival prospects and ultimately reduce welfare. The effects are qualitatively identical to the capital cost subsidy.

**PROPOSITION 5 (Wage Tax).** (a) *A higher wage tax raises the interest rate and expands market size.* (b) *The tax expands entrepreneurship and the number of mature firms, but erodes firm value.* (c) *It impairs incentives for effort and advice and reduces the survival probability.* (d) *Introducing a small wage tax leads to a first order welfare loss.*

As a corollary, a *subsidy* to wage income would restrict entry, leading to fewer firms with higher values. The subsidy could raise welfare since increased firm values sharpen incentives for joint effort. The start-up investment tax in the preceding subsection and the wage subsidy in this subsection can be compared to De Meza and Webb (1987) who argue, for entirely different reasons, that entrepreneurial entry should be discouraged.

#### 4. CONCLUSIONS

This chapter has proposed an equilibrium model of the venture capital industry and has investigated the consequences and appropriateness of fiscal policy for the quality and quantity of venture capital financed entrepreneurship. Such an analysis is important for several reasons. First, the creation of young entrepreneurial firms is a significant factor in promoting employment and innovation in a growing economy. Second, venture capital has become an increasingly important source of finance for start-up firms over the last decades in virtually all industrial countries. In combining financing of new firms with active advice and networking support, venture capital can help the professionalization of their portfolio companies and add value to the investments. For this reason, venture capital backed firms appear to outperform similar firms without access to venture capital, making them a particularly important source of job growth and innovation in the economy. Empirical research for the U.S. such as Kortum and Lerner (2000) has shown indeed that a disproportionately large share of industrial innovation in the U.S. originates with venture capital backed firms. Policy makers are thus much concerned about creating the right policy environment for a dynamic venture capital industry.

Third, the business community at large as well as the venture capital industry itself have repeatedly questioned whether existing public policies are sufficiently conducive to the development of start-ups firms. For instance, the European Venture Capital and Private Equity Association has twice issued a benchmarking report on the conditions for entrepreneurship in its member countries. These reports define a favorable environment by low corporate income taxes and taxes on capital gains on individual investments in entrepreneurial firms. They also point to the importance of fiscal subsidies to research and development and other early stage investment cost. It is therefore important to scrutinize the economic rationale as well as the effectiveness of these policies in stimulating the venture capital sector.

Rather than simply arguing for high subsidies and low taxes to stimulate entrepreneurship, as is often done, a stringent theoretical framework is called for in order to appropriately assess the role of relevant taxes and subsidies in determining the level and quality of venture capital backed entrepreneurship and economic efficiency. We have proposed a structural equilibrium model of the venture capital industry that emphasizes the need for outside risk capital and points to the importance of incentive problems that entrepreneurs and financiers may face in a typical, innovative start-up company. With this formal framework at hand, we have derived some important policy implications.

Our results imply that the taxation of capital gains derived from young firms may be quite harmful to the quality of venture capital financed entrepreneurship and may diminish welfare. Further, and perhaps surprisingly,

corporate taxes are not only harmful to the expansion investment and value of mature firms but could be equally harmful to start-up firms which have not yet begun to actually pay the tax. In reducing mature firm value to be realized at the end of the start-up phase, the corporate tax impairs the incentives of entrepreneurs and venture capitalists for effort and active advice at the early stage of the firm's development. It may therefore contribute to an overly high failure rate and harm the quality of venture capital backed firms. Our analysis thus lends some support to the advocates of cutting the capital gains tax or giving corporate tax relief to small innovative firms. However, such tax relief should be confined to venture capital backed firms only. In terms of practical tax policy, two issues should be noted. First, the burden of the capital gains tax may already be quite low compared to other capital income taxes. The deferral of capital gains until realization implies interest gains to the tax payer that much reduce the actual tax burden. Second, there might be some practical difficulties in selectively applying a tax break to venture capital backed firms only.

Most of the real world programs to stimulate business creation involve a subsidy to the cost of capital in one or the other form. However, since these subsidies are given early on and are not success-related, they are not useful for sharpening incentives for effort and advice. Because they boost the rate of business creation, they may actually reduce equilibrium venture returns and thereby discourage effort and advice within VC-backed start-ups. When reducing the quality of entrepreneurship this way, investment subsidies may turn out to be quite undesirable.

Our insights as to the role of taxes and subsidies show that they can be combined in a self-financing way to improve the quality of venture capital investments. Instead of a subsidy, a tax on start-up capital cost is proposed with the revenue used to finance a selective tax cut on the capital gains derived from venture capital backed investment. This package replaces a non-performance related subsidy with a success related tax cut, sharpens incentives within start-up firms and should thereby contribute to a more active style of venture capital financing. Very importantly, this package implies a zero net tax or subsidy per project and thereby does not come at the expense of the general tax payer.

There are, of course, other arguments such as the possibility of new innovative firms creating spillovers to other firms which may be relevant in designing an appropriate policy vis-à-vis high-technology start-ups. Such technological spillovers might call for a net subsidy to the sector (see Keuschnigg, 2003). However, even in this case our analysis implies that the subsidy should be given in form of a selective tax break which strengthens incentives by rewarding success.

As always, the results of policy analysis must be seen in light of the specifics of the theoretical model used. We have chosen to focus on what we believe are important incentive (moral hazard) problems related to

value generation in start-up firms. Other information problems (of the adverse selection type) may be relevant as well in designing an appropriate policy vis-à-vis high-technology start-ups. This literature, as summarized by De Meza (2002), among others, partly argues for a tax on entry such as an interest tax on capital investment to offset a tendency for entrepreneurial overinvestment. Quite reassuringly, there is no conflict with our analysis. Our framework does not call for a net subsidy to encourage entry but only recommends a restructuring of existing taxes and subsidies to strengthen incentives and thereby raise survival prospects of start-up firms. In this sense, our analysis argues not for more, but for more successful venture capital backed firms. Hopefully, these issues will be further addressed in future research within a unified framework that considers both the screening and value added activities of venture capitalists.

## NOTES

<sup>1</sup> See EVCA Barometer June 2005 on <http://www.evca.com>.

<sup>2</sup> See also the related Press Release from EVCA of May 24, 2004.

<sup>3</sup> We discuss the robustness of this conclusion in Section 2.6.

<sup>4</sup> For simplicity, we assume complete loss-offset in capital gains taxation. The consequence of incomplete loss-offset is analyzed by Keuschnigg and Nielsen (2003b).

<sup>5</sup> For simplicity, we assume away any other cost of VC entry, thus making VC financing perfectly competitive and allowing no more than zero profits. Keuschnigg (2003) considers search costs of VCs and entrepreneurs in the seed phase prior to entering a financing relationship. To cover these costs, both agents must earn strictly positive rents during the start-up phase. Kanniainen and Keuschnigg (2004) assume fixed entry costs of VC firms. This allows to contrast a short-run equilibrium with a fixed number of VCs and positive rents with a long-run equilibrium with free entry and zero profits (see the graphical analysis of Gompers and Lerner, 2002, with a similar distinction of short- and long-run equilibrium). These extensions do not affect our basic conclusions. We thus focus only on the perfectly competitive, long-run equilibrium in this chapter.

<sup>6</sup> See Keuschnigg and Nielsen (2004a) for this result. There we show that the privately optimal equity share  $s$  changes only when there are differential capital gains taxes on entrepreneurs and VCs, or when the tax treatment of the upfront payment  $B$  is different from the capital gains tax.

<sup>7</sup> This symmetry assumption excludes problems of adverse selection and helps to focus on the moral hazard during the start-up phase.

<sup>8</sup> In the discussion paper version (Keuschnigg and Nielsen, 2004c), the entry decision is reconciled with consumption and savings choice. A simple specification of preferences implies that differences in life-time welfare among alternative occupations are uniquely related to differences in income, adjusted for effort.

<sup>9</sup> Poterba (1989a, b) and Gompers and Lerner (1998, 2002) introduced a graphical supply and demand analysis, albeit without an explicit structural model. In Gompers and Lerner (2002), the expected return on VC investments adjusts to establish equilibrium in the VC industry by reconciling the willingness of investors to supply capital with the number of entrepreneurial firms meeting the return requirement (see their Figure 9-5). In contrast to these authors, we assume the supply of financial capital to be perfectly elastic and thus do not consider it as a bottleneck of the

VC industry. In our model, the number of firms demanded stems from price elastic demand for the output of VC backed firms.

<sup>10</sup> If we had combined the model with horizontal product differentiation in the output market as in Keuschnigg (2003), there would also be a reason to additionally encourage entry to enlarge product variety. In this case, entry subsidies would help to internalize the technological spillovers from launching new products. See the analysis of Keuschnigg (2004b) within an endogenous growth framework.

<sup>11</sup> Such incentive problems in teams have been first analyzed by Holmstrom (1982) and were applied, among others, by Aghion and Tirole (1994) to analyze output of innovation teams.

<sup>12</sup> The corporate tax reduces entrepreneurship if the tax is neutral with respect to expansion investment ( $\theta = 1$ ), or if the interest elasticity of (second period) output demand  $\sigma_0$  exceeds the elasticity  $\eta$  of capital demand per firm with respect to the user cost,  $\sigma_0 \geq \eta$ . In Figure 9-4, a large  $\sigma_0$  implies that any given interest increase triggers a large increase in market size, leading to a steep slope of the demand schedule. A small capital demand elasticity  $\eta$  leads to a relatively smaller upward shift of the demand curve. It can thus be illustrated graphically that this condition works to erode entrepreneurship.

<sup>13</sup> Assuming a fixed interest as in a small open economy would not change the qualitative results of propositions 1 to 3 which do not hinge on the general equilibrium effects on the interest rate.

<sup>14</sup> For a more formal exposition of the effects of the self-financing policy we refer to Keuschnigg and Nielsen (2004a).

<sup>15</sup> Note that the policy would work even better in an open economy where any adjustment in the interest rate and mature firm value is limited.

## REFERENCES

- Aghion, P. and Tirole, J. (1994). The Management of Innovation. *Quarterly Journal of Economics* 109: 1185–1209.
- Audretsch, D. B. and Lehmann, E. (2002). *Debt or Equity? The Role of Venture Capital in Financing the New Economy in Germany*. CEPR Discussion Paper No. 3656, Centre for Economic Policy Research, London.
- Black, B. S. and Gilson, R. J. (1998). Venture Capital and the Structure of Capital Markets: Banks versus Stock Markets. *Journal of Financial Economics* 47: 243–277.
- Boadway, R. and Tremblay, J. F. (2005). Public Economics and Start-up Entrepreneurship. In *Venture Capital, Entrepreneurship and Public Policy*, V. Kannianen and C. Keuschnigg (Eds), pp. 198–219. MIT Press, Cambridge, MA.
- Bottazzi, L. and Da Rin, M. (2002). Venture Capital in Europe and the Financing of Innovative Companies. *Economic Policy* 34: 231–269.
- Casamatta, C. (2003). Financing and Advising: Optimal Financial Contracts with Venture Capitalists. *Journal of Finance* 58: 2058–2085.
- Cornelli, F. and Yosha, O. (2003). Stage Financing and the Role of Convertible Securities. *Review of Economic Studies* 70: 1–32.
- Cressy, R. (2002). Introduction: Funding Gaps. *Economic Journal* 112: F1–F16.
- Cullen, J. B. and Gordon, R. H. (2002). *Taxes and Entrepreneurial Activity: Theory and Evidence for the U.S.* NBER Working Paper No. 9015, National Bureau of Economic Research, Cambridge, MA.
- Da Rin, M., Nicodano, G. and Sembenelli, A. (2005). *Public Policy and the Creation of Active Venture Capital Markets*. Unpublished Discussion Paper, Universita di Torino, Italy.

- De Meza, D. and Webb, D. C. (1987). Too Much Investment: A Problem of Asymmetric Information. *Quarterly Journal of Economics* 102: 281–292.
- De Meza, D. (2002). Overlending? *Economic Journal* 112: F17–F31.
- EVCA (2004). *Benchmarking European Tax & Legal Environments*. Report, European Private Equity and Venture Capital Association.
- Fuest, C., Huber, B. and Nielsen, S. B. (2003). Why Is the Corporate Tax Rate Lower Than the Personal Tax Rate? The Role of New Firms. *Journal of Public Economics* 87: 157–174.
- Gentry, W. M. and Hubbard, R. G. (2000). Tax Policy and Entry into Entrepreneurship. *American Economic Review* 90: 283–287.
- Gompers, P. A. and Lerner, J. P. (1996). The Use of Covenants: An Empirical Analysis of Venture Partnership Agreements. *Journal of Law and Economics* 39: 463–498.
- Gompers, P. A. and Lerner, J. P. (1998). What Drives Venture Capital Fundraising? *Brookings Papers on Economic Activity Microeconomics* 149–192.
- Gompers, P. A. and Lerner, J. (1999). *The Venture Capital Cycle*. MIT Press, Cambridge, MA.
- Gompers, P. A. and Lerner, J. (2001). The Venture Capital Revolution. *Journal of Economic Perspectives* 15: 145–168.
- Gompers, P. A. and Lerner, J. (2002). Short-Term America Revisited? Boom and Bust in the Venture Capital Industry and the Impact on Innovation. *Innovation Policy and the Economy* 3: 1–27.
- Gordon, R. (1998). Can High Personal Tax Rates Encourage Entrepreneurial Activity? *IMF Staff Papers* 45: 49–80.
- Hart, O. (2001). Financial Contracting. *Journal of Economic Literature* 39: 1079–1100.
- Hege, U., Palomino, F. and Schwienbacher, A. (2003). *Determinants of Venture Capital Performance: Europe and the United States*. RICAPE Working Paper No. 001, Financial Markets Group, London School of Economics.
- Hellmann, T. (1998). The Allocation of Control Rights in Venture Capital Contracts. *RAND Journal of Economics* 29: 57–76.
- Hellmann, T. and Puri, M. (2000). The Interaction Between Product Market and Financing Strategy: The Role of Venture Capital. *Review of Financial Studies* 13: 995–984.
- Hellmann, T. and Puri, M. (2002). Venture Capital and the Professionalization of Start-Up Firms. *Journal of Finance* 57: 169–197.
- Holmstrom, B. (1982). Moral Hazard in Teams. *Bell Journal of Economics* 13: 324–340.
- Inderst, R. and Mueller, H. M. (2004). The Effect of Capital Market Characteristics on the Value of Start-Up Firms. *Journal of Financial Economics* 72: 319–356.
- Kannianen, V. and Keuschnigg, C. (2004). Start-up Investment With Scarce Venture Capital Support. *Journal of Banking and Finance* 28: 1935–1959.
- Kaplan, S. N. and Stromberg, P. (2001). Venture Capitalists as Principals: Contracting, Screening, and Monitoring. *American Economic Review* 91: 426–430.
- Kaplan, S. N. and Stromberg, P. (2003). Financial Contracting Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies* 70: 281–315.
- Kaplan, S. N. and Stromberg, P. (2004). Characteristics, Contracts, and Actions: Evidence From Venture Capital Analysis. *Journal of Finance* 59: 2173–2206.
- Keuschnigg, C. (2003). *Optimal Public Policy for Venture Capital Backed Innovation*. CEPR Discussion Paper No. 3850, Centre for Economic Policy Research, London.
- Keuschnigg, C. (2004a). Taxation of a Venture Capitalist With a Portfolio of Firms. *Oxford Economic Papers* 56: 1–22.
- Keuschnigg, C. (2004b). Venture Capital Backed Growth. *Journal of Economic Growth* 9: 239–261.
- Keuschnigg, C. and Nielsen, S. B. (2003a). Tax Policy, Venture Capital, and Entrepreneurship. *Journal of Public Economics* 87: 175–203.



- Keuschnigg, C. and Nielsen, S. B. (2003b). Taxes and Venture Capital Support. *European Finance Review* 7: 515–539.
- Keuschnigg, C. and Nielsen, S. B. (2004a). Start-ups, Venture Capitalists, and the Capital Gains Tax. *Journal of Public Economics* 88: 1011–1042.
- Keuschnigg, C. and Nielsen, S. B. (2004b). Taxation and Venture Capital Backed Entrepreneurship. *International Tax and Public Finance* 11: 369–390.
- Keuschnigg, C. and Nielsen, S. B. (2004c). *Public Policy, Start-up Entrepreneurship, and the Market for Venture Capital*. University of St Gallen Working Paper, Switzerland: <http://www.iff.unisg.ch>, page Research/DP Public Finance.
- Kortum, S. and Lerner, J. (2000). Assessing the Contribution of Venture Capital to Innovation. *RAND Journal of Economics* 31: 674–692.
- Lerner, J. (1994). The Syndication of Venture Capital Investments. *Financial Management* 23: 16–27.
- Lerner, J. (1999). The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program. *Journal of Business* 72: 285–318.
- Micchelacci, C. and Suarez, J. (2004). Business Creation and the Stock Market. *Review of Economic Studies* 71: 459–481.
- Parker, S. C. (2003). Asymmetric Information, Occupational Choice and Government Policy. *Economic Journal* 113: 861–882.
- Poterba, J. M. (1989a). Capital Gains Tax Policy toward Entrepreneurship. *National Tax Journal* 42: 375–389.
- Poterba, J. M. (1989b). Venture Capital and Capital Gains Taxation. In *Tax Policy and the Economy*, Vol. 3, L. H. Summers (Ed.), pp. 47–67. MIT Press, Cambridge, MA.
- Repullo, R. and Suarez, J. (2004). Venture Capital Finance: A Security Design Approach. *Review of Finance* 8: 75–108.
- Rosen, H. S. (2005). Entrepreneurship and Taxation: Empirical Evidence. In *Venture Capital, Entrepreneurship and Public Policy*, V. Kannianen and C. Keuschnigg (Eds), pp. 251–279. MIT Press, Cambridge, MA.
- Schmidt, K. M. (2003). Convertible Securities and Venture Capital Finance. *Journal of Finance* 43: 1139–1166.
- Sorensen, M. (2005). *How Smart is Smart Money? An Empirical Two-Sided Matching Model of Venture Capital*. Unpublished paper, University of Chicago.
- Stiglitz, J. E. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review* 71: 393–410.
- Ueda, M. (2004). Bank Versus Venture Capital. Project Evaluation, Screening, and Expropriation. *Journal of Finance* 59: 601–621.
- Wallsten, S. J. (2000). The Effects of Government-Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research Program. *RAND Journal of Economics* 31: 82–100.

COLIN M. MASON

*Hunter Centre for Entrepreneurship, University of Strathclyde, Glasgow G1 1XH, United Kingdom*

*E-mail: colin.mason@strath.ac.uk Q*

## 10. Informal Sources of Venture Finance

### 1. INTRODUCTION

One of the most interesting insights from the Global Enterprise Monitor (GEM)—an annual international survey of entrepreneurial activity in some 40 countries—is that informal sources of finance overwhelm formal ones (Bygrave et al., 2003). GEM’s methodology captures two sources of informal financing: family members (often termed “love money”) and other individuals, the latter comprising investors who have come to be known as *business angels*<sup>1</sup> who invest in new and young businesses where there is no family connection. Some 3.4% of the adult population in the 18 countries where information is available meet the definition of being an informal investor.<sup>2</sup> They provide \$196m per year to new and growing companies, equivalent to 1.1% of the GDP of these countries, and account for between 60% and 90% of total venture capital, including institutional sources. In the U.S., 5% of the population are informal investors, collectively investing \$108b per annum, which is 3.5 times the amount invested by venture capital funds in seed and start-up investments (Bygrave and Reynolds, 2004). Most informal investment flows to family members and friends. In the 18 GEM countries, some 50% of informal investment goes to relatives, 29% to friends/neighbors, 11% to work colleagues and just 8% to a stranger (Bygrave et al., 2003). The proportions for the U.S. are almost identical (Bygrave and Reynolds, 2004).

This chapter focuses on business angels who collectively make up what is termed the *informal venture capital market* (in contrast to the formal, or institutional venture capital market—see Chapter 9). The role of the family in funding entrepreneurial ventures is an underresearched topic. However, by definition, access to finance from family members is constrained by ties of

blood and marriage, and is therefore only available to other family members. Accordingly, it does not constitute a market. If an entrepreneur is unfortunate enough to come from an impoverished family then this source of potential funding is closed off. Business angels, in contrast, do invest in businesses that are owned by strangers (as well as those owned by acquaintances) and it is quite appropriate for an entrepreneur seeking finance to approach them for funding. Indeed, business angels constitute the largest pool of equity capital available for start-up and emerging companies in advanced economies (Gaston, 1989a). Moreover, as we will see later in this chapter, they contribute much more than just money to their investee companies.

The research base on informal venture capital is limited. Prowse (1998, p. 786) commented in the late 1990s that “the angel market operates in almost total obscurity. Very little is known about its size, scope, the type of firms that raise angel capital and the types of individuals that provide it.” Indeed, prior to the 1980s, business angels were unknown both to researchers and policy makers. The emergence of informal venture capital as a distinct topic within the entrepreneurship literature is therefore relatively recent. Pioneering studies by Wetzel (1981, 1983, 1987) and others (e.g., Tymes and Krasner, 1983; Haar et al., 1988; Gaston, 1989b) began to establish its importance in the U.S. during the 1980s. Subsequent studies in Canada (Riding and Short, 1989; Short and Riding, 1989), Europe (e.g., Harrison and Mason, 1992; Landström, 1993; Mason and Harrison, 1994; Reitan and Sørheim, 2000; Brettel, 2003; Stedler and Peters, 2003; Paul et al., 2003), Australasia (Hindle and Wenban, 1999; Infometrics, 2004) and Asia (Tashiro, 1999; Hindle and Lee, 2002) confirmed that business angels were not just a U.S. phenomenon.<sup>3</sup> There is also evidence of angel activity in less developed regions of the world, such as South America (Pereiro, 2001).<sup>4</sup> While these studies are in what has been termed the “ABC” tradition of angel research (attitudes, behaviors and characteristics), other so-called second-generation studies have sought to develop a more in-depth understanding of business angel investment activity by focusing on process (Mason and Harrison, 2000a). Accordingly, Prowse’s (1998) claim is no longer valid. Nevertheless, both the body of literature on informal venture capital and the number of active researchers remain small in comparison to the amount of scholarly activity that is devoted to investigating institutional venture capital. This is despite the much greater role of informal venture capital in funding the start-up and initial growth of entrepreneurial ventures.<sup>5</sup> A further limitation is that much of the research lacks theoretical foundations and simply reports survey findings. Only by attracting more scholarly effort and producing research that is theoretically informed will informal venture capital lose this Cinderella status.

This chapter adopts a supply-side perspective. This reflects the focus of the overwhelming majority of the literature on this topic which is largely

concerned with who business angels are, what motives them and how they invest. By way of contrast, there is very little literature that takes a demand-side perspective, looking at informal venture capital from the viewpoint of the entrepreneur. The chapter is structured as follows. The next section provides an overview of the ‘ABC’ of angels. Section 3 reviews the economic importance of business angels, highlighting the nature of their investments and the size of the market. This is followed in Section 4 by a brief discussion of government efforts to expand the supply of informal venture capital. Section 5 adopts an investment process perspective. Section 6 examines the emergence of new organizational formats for angel investing. The concluding section offers some thoughts on future research possibilities.

## 2. THE INFORMAL VENTURE CAPITAL MARKET: AN OVERVIEW

The informal venture capital market comprises business angels who are conventionally defined as *high net worth individuals who invest their own money, along with their time and expertise, directly in unquoted companies in which they have no family connection, in the hope of financial gain*. Several aspects of this definition need to be highlighted in order to emphasize the distinctiveness of business angels as a type of investor.

*High net worth.* Having wealth is a prerequisite for becoming a business angel. Business angels invest upward of £10,000 per deal (sometimes in excess of £100,000) and typically have a portfolio of two to five investments (some angels have more). However, they are not investing their entire savings in this way. Because of the high risk of investing in unquoted companies, most angels allocate just 5%–15% of their overall investment portfolio to such investments. Thus, if these investments fail, as they often do, the losses will not affect their lifestyle. Some rather dated evidence on the wealth of angels suggests that they tend to be “comfortably” off rather than super-rich. Gaston (1989b) reported that one in three business angels in the U.S. had a net worth (excluding principal residence) in excess of \$1 million. There were also “very few” millionaires in the study by Haar et al. (1988). Mason and Harrison (1994) noted that only 19% of U.K. business angels were millionaires.

*Investing their own money.* The fact that angels are investing their own money distinguishes them from institutional venture capital funds whose investment funds come from such sources as pension funds, banks and foundations and, as a result, have a legal duty of care for how they invest such funds. First, business angels do not have to invest if they do not find appropriate investments, whereas venture capital funds have a fixed life, typically ten years, over which the fund must invest and exit. Second, they can make quicker investment decisions (Freear et al., 1995). Third, angels have less need for

specialist financial and legal due diligence, so the costs for the investee business are lower. Fourth, business angels can adopt idiosyncratic investment criteria whereas venture capital funds have raised their investment funds to invest in specific types of businesses and so must follow these investment criteria when investing.

*Direct.* Business angels make their own investment decisions as opposed to investing in some form of pooled investment vehicle in which the investment decisions are made by fund managers. This implies that those people who become business angels have both the personal networks that will provide a flow of investment opportunities and the competence to undertake the appraisal of new and young entrepreneurial companies. Indeed, a consistent theme in the literature is that the majority of business angels are successful, cashed-out entrepreneurs, while the remainder either have senior management experience in large businesses or have specialist business expertise (e.g., accountant). On account of these backgrounds, such individuals have access to deal flow and the competence to make investment decisions. Becoming a business angel is therefore a way for such individuals to recapture their successful experience, making investments based on the analytical skills and intuition that they have developed in business. This reinforces their self image and sustains recognition in the communities in which they live (Margulis and Benjamin, 2000). However, it is fair to say that competence levels among business angels is variable, and a career as a successful entrepreneur, or in a senior position in a large company, does not necessarily provide an individual with all of the skills required to be a successful business angel. Angels report in surveys (Sørheim, 2003) that their initial investments involved a steep learning curve.

*Time and expertise.* Part of the investment approach of business angels involves the support their investee businesses through a variety of hands-on roles, including mentoring, the provision of strategic advice, networking and in some cases direct involvement in a specific functional capacity. This has prompted the description of informal venture capital as being “capital and consulting.” The opportunity to be involved with a business start-up is a significant motive for business angels. Involvement also reduces information asymmetries and moral hazard and so is a means of risk reduction.

*Unquoted companies.* Business angels are investing in unquoted companies as opposed to companies that are listed on a stock market. As we will see more clearly in the next section, while angels invest in all sorts of situations, including management buyouts and buy-ins and rescue/turnaround situations, their typical investment is in a new or recently started business. The key point here is that business angels want to be active investors in the companies in which they invest, helping them to grow, whereas stock market investing is passive.

*Financial gain.* Business angels are investing in the hope of achieving a financial return, typically in the form of a capital gain that is accomplished

through some form of harvest event such as an acquisition of the investee company or an IPO. However, psychic income is also an important motivation. Studies are consistent in identifying that the fun and enjoyment that is derived from such investments is an important subsidiary reason for becoming a business angel. This links back to an earlier point: business angels are also characterized as being *hands-on investors*. The ability to provide support to investee companies reinforces the tendency for business angels to have a business background. Some angels also express altruistic motives. Paul et al. (2003) quoted one Scottish angel as follows: “don’t get me wrong, I want to make money. But I’ve done well out of Scotland and I’d like to help others to do the same.” U.S. evidence indicates that most business angels would be willing to forego *some* financial return either to invest in businesses that were seen as socially beneficial (Sullivan, 1994) or simply to support new entrepreneurs (Wetzel, 1981). Evidence of altruistic motives is much weaker in other countries.

One of the striking features in the literature is the remarkable consistency in the characteristics of business angels across countries. Japan is the only country where research suggests that angels have a distinctively different profile (Tashiro, 1999). The profile of the *typical* business angel is characterized as follows:

- *Male*. Studies in various countries are consistent in finding that upward of 95% of business angels are male. This can be attributed to the relatively small numbers of women who have built successful entrepreneurial companies or hold senior positions in large companies. However, the small minority of women who are business angels have similar characteristics to those of their male counterparts (Harrison and Mason, 2005).
- *In the 45–65 year age group*. This reflects the length of time required to build significant personal net worth, the greater amount of discretionary wealth of this age group as their children cease to become financially dependent on them, and the age at which people with a successful business career might chose, or be forced to, disengage. Becoming a business angel is often a way in which such individuals to remain economically active. For example, cashed out entrepreneurs in their 40s or 50s often report that they became business angels because they quickly became bored by a life of leisure—as one angel noted, “the attractions of playing golf seven days a week quickly palls.” There are some international differences. Angels are slightly younger in the U.S. and slightly older in Nordic countries (Landström, 1993). Recent studies hint that business angels may be becoming slightly younger (e.g., Infometrics, 2004). This may be linked, at least in part, to the acquisition frenzy of the closing

years of the technology boom of the late 1990s which enabled a lot of younger entrepreneurs to cash-out.

- *Successful cashed-out entrepreneurs.* Most business angels have had experience of business start-up and growth. As Freear et al. (1992, p. 379) noted, this implies that many angels “have acquired the kind of experience...that it takes to start, manage and harvest a successful entrepreneurial venture. In a sense their entire professional careers have prepared them to conduct the due diligence necessary to evaluate the merits and risks of prospective investments and to add value of their know how to the ventures they bankroll.” The remainder are typically either people who have held senior positions in large companies or have specialist commercial skills and are involved in working with entrepreneurial companies (e.g., accountants, consultants, lawyers) and whose wealth is derived from high income. It is also important to emphasize that non-business professionals (e.g., doctors, dentists) and public sector employees are conspicuous by their absence from the ranks of business angels (Gaston, 1989b).
- *Well educated.* Economic success is underpinned by a high level of education. Business angels typically have a university degree and/or professional qualifications. However, angels with PhDs are rare. This reflects other research that suggests that the relationship between education and entrepreneurship is an inverted U-shape (i.e., both too little and too much education is a hindrance to entrepreneurial behaviors) (Reynolds, 1997).

There have been surprisingly few attempts to compare business angels with noninvestors. Lindsay (2004) found that angels score more highly on measures of entrepreneurial orientation—pro-activeness, innovativeness, risk-taking strategies—which, in turn, suggests that they act in an entrepreneurial manner in undertaking their investment activities. However, this might simply reflect the entrepreneurial background of most business angels. Duxbury et al. (1996) suggested that angels are distinctive from non-investors in terms of their psychological traits, with an internal locus of control, very high need for achievement (nAch), a moderately high need for affiliation and autonomy and are intrinsically motivated. But here again these are also entrepreneurial traits.

This profile masks considerable heterogeneity in the business angel population, not so much in terms of their demographics, but rather in their motivation and investment focus. The most basic distinction is between *active angels*—those individuals with experience of investing and who are continuing to look for investments, *latent angels*—inactive investors who have made investments in the past, and *virgin angels*—individuals who are looking to invest but have yet to make their first investment (Coveny and Moore, 1998).

TABLE 10-1 Differences between types of business angel: I—Coveney and Moore

| Characteristic                     | Entrepreneur angel | Wealth maximizing | Income seeking |
|------------------------------------|--------------------|-------------------|----------------|
| Total funds invested               | £590,000           | £131,000          | £35,000        |
| No of investments                  | 3.4                | 2.1               | 1.5            |
| Personal net worth                 | 74% > £1m          | 43% > £1m         | 75% > £1m      |
| Reason for investing               | returns/fun        | returns           | job/income     |
| <i>Typical deal</i>                |                    |                   |                |
| Average total amount invested      | £174,000           | £54,000           | £24,000        |
| Average initial amount invested    | £111,000           | £21,000           | £17,000        |
| Average number of rounds           | 2                  | 1.75              | 1.5            |
| Average number of co-investors     | 2.3                | 2.5               | 3.0            |
| Average size of equity stake taken | 38%                | 31%               | 20%            |

Note: based on a survey of “nearly 500 business angel investors/potential investors ... and 467 actual investment deals involving a total level of funds of more than £50 million” (Coveney and Moore, 1998: 8). However, the methodology for classifying investors is not explained.

There are several classifications of active investors. Gaston (1989b) identified ten distinct types of business angel but without elaborating on the methodological basis for the classification. Coveney and Moore (1998) identified three types of business angel based on their level of entrepreneurial activity and intensity of investment activity (see Table 10-1):

- *Entrepreneur angels*: the most active in terms of number of investments and amount invested, the most experienced angels and also the most wealthy. Their preference is to invest at start-up and enjoyment is a major motivation. Their key investment criterion is the personality of the entrepreneur. Entrepreneur angels are also the most open to investing outside of their own field of experience. They are unlikely to play a role in the day-to-day management of their investee companies.
- *Income seeking angels*: significantly less wealthy investors, less active and less motivated by fun and enjoyment considerations, tend to invest in industries in which they are familiar and looking for a formal management role in the ventures which they finance.
- *Wealth maximizing angels*: predominantly self-made investors but includes some with inherited wealth, interested primarily in the financial return, more likely to invest in industries in which they have personal experience and more likely to take a full-time position in their investee businesses.

Sørheim and Landström (2001) used cluster analysis to differentiate Norwegian business angels in terms of their competence and investment activity. This produces four distinct types of business angel (Table 10-2):



TABLE 10-2 *Differences between types of business angel: II—Sørheim and Landström*

|  | Lotto investors | Traders | Analytical investors | Business angels |
|--|-----------------|---------|----------------------|-----------------|
| Investors with gross income over 500,000 NOR (%)     | 14              | 42      | 39                   | 77              |
| Net worth over 2 million NOR (%)                     | 6               | 29      | 17                   | 74              |
| Number of investment proposals                       | 8.2             | 19.9    | 7.9                  | 44.5            |
| Number of investments made                           | 1.4             | 4.5     | 1.7                  | 7.3             |
| Invested with other business angels (%)              | 31              | 48      | 59                   | 83              |
| Invested with banks, venture capital funds, etc. (%) | 12              | 25      | 18                   | 43              |
| Functioned as lead investor                          | 2               | 11      | 5                    | 42              |
| Served as board member for investee businesses       | 3               | 14      | 34                   | 61              |
| Acted as consultant to investee businesses           | 2               | 3       | 8                    | 24              |

Based on a sample of 425 “informal investors.”

- *Lotto investors* (30%): low investment activity level and limited experience of starting and running businesses. They make very few investments and have limited ability to add value to their investments.
- *Traders* (24%): high investment activity but limited experience of starting and running entrepreneurial businesses. They are keen to invest but have limited ability to add value.
- *Analytical investors* (21%): low level of investment activity but possess fairly high competence.
- *Business angels* (25%): very high level of investment activity and high competence.

From a demand-side perspective these studies underline the differentiated nature of the supply of informal venture capital. Clearly, “not everybody’s money is green.” The implication for entrepreneurs is that they must ensure that the type of business angel who is offering to invest is both willing and capable of contributing the value-added that they require.

Other studies have focused on specific types of business angel. Kelly and Hay (1996, 2000) focused on the most active investors who account for a disproportionate amount of investment activity. They note that such angels are more financially driven and formalized in their approach, which they suggest reflects their experience of living through unforeseen problems and obstacles. Visser and Williams (2001) examined “takeover and turnaround artists”—business angels who specialize in investing in distressed companies with the aim of turning them around to start on a growth path again.<sup>6</sup> As they note, these investors are “performing the same function as . . . other types of business angels. . . —breathing new life into a business—but at the other end of the

business spectrum—when the business is about to die” (Visser and Williams, 2001, p. 2).

### 3. THE ECONOMIC SIGNIFICANCE OF THE INFORMAL VENTURE CAPITAL MARKET

The informal venture capital market is recognized as playing a vital role in economic development at both national and local/regional scales. Indeed, one U.K. government report argued that “an active informal venture capital market is a pre-requisite for a vigorous enterprise economy. . .” (ACOST, 1990, p. 41). There are three aspects of the informal venture capital market which are significant from an economic development perspective.

First, the amount of finance that business angels have invested, or have available to invest, is significant. Unfortunately, it is impossible to be precise about the number of business angels, the number of investments made and the amount invested. This is because there is no obligation for business angels to identify themselves or register their investments. Indeed, the vast majority of business angels strive to preserve their anonymity and are secretive about their investment activity, not least to avoid being inundated by entrepreneurs and other individuals seeking to persuade them to invest or provide financial support for other causes (Benjamin and Margulis, 2000). Thus, all measures of the size of the informal venture capital market are fairly crude estimates. Gaston (1989a) estimated that in the U.S. business angels invest 13 times more dollars than venture capital funds and make 40 times more investments. A more up-to-date estimate by Sohl (2003) suggested that there are 300,000–350,000 business angels in the U.S., investing approximately \$30 billion per annum in close to 50,000 ventures. Venture capital funds, in contrast, invest \$30–\$35 billion in fewer than 3,000 entrepreneurial ventures. The equivalent estimate for the U.K. is 20,000–40,000 business angels investing £0.5 billion—£1 billion per annum in 3,000–6,000 companies. They make eight times as many investments in start-up companies as venture capital funds (Mason and Harrison, 2000b). However, these calculations of the amounts invested by business angels are an underestimate of the size of the informal venture capital market. First, most business angels have further funds available to invest (Coveney and Moore, 1998; Mason and Harrison, 1994, 2002a) but cannot identify appropriate investment opportunities. This uncommitted capital is substantial: one study reported that it exceeded the amount invested by the respondents in the three years prior to the survey (Mason and Harrison, 2002a). Second, there is a substantial pool of potential, or virgin, business angels who share the characteristics of active angels but have not entered the market (Freear et al., 1994a; Coveney and Moore, 1998). However, with appropriate forms

of support—such as help with deal flow and with the technical aspects of investing—they could be encouraged to enter the market (Mason and Harrison, 1993; Freear et al., 1994a). Sohl (1999) estimated that these potential angels exceed the number of active investors by a factor of five to one.

The economic significance stems from where this capital is invested. Finance from business angels occupies a crucial place in the spectrum of finance available to growing businesses. In terms of *size of investment*, business angels invest in what is often termed (at least in Europe) the “equity gap,” providing amounts of finance that are beyond the ability of entrepreneurs to raise from their own resources and from family and below the minimum investment threshold of venture capital funds<sup>7</sup>—a figure that is in excess of £1m in the U.K. and \$5m in the U.S. (Sohl, 2003). Business angels, investing on their own or in small *ad hoc* groups, will typically invest up to £100,000, or even £250,000, while the larger angel syndicates (see Section 6) will make investments of £500,000 and above. This is usually provided in the form of equity or a combination of equity plus loans. However, all-loan investments are by no means unusual.<sup>8</sup> In terms of *stage of business development*, investments by business angels are skewed toward the seed, start-up and early growth stages whereas venture capital funds focus on later stage deals. The role of business angels in seeding new ventures has become even more critical in recent years as institutional venture capital funds in North America and Europe have raised their minimum investment size and continued to shift their investment focus to later stage investments (Jensen, 2002; Sohl, 2003).

The second factor which underpins the economic significance of the informal venture capital market is the hands-on involvement of business angels in their investee businesses. Demand-side studies indicate that many entrepreneurs are seeking “smart money” and for this reason business angels are valued ahead of other funding sources (Cressy and Olofsson, 1997; Lindström and Olofsson, 2001; Sætre, 2003). It has already been noted that business angels derive considerable psychic income from this involvement. Their entrepreneurial and business backgrounds have also been highlighted. Further discussion of the nature of this involvement can be found in Section 5: suffice to say at this point that it ranges from informal coaching, mentoring and advice to Board participation. Business angels typically invest in industries and markets with which they are familiar. As a consequence, the entrepreneurs who are funded by business angels derive consider value from the expertise, knowledge and experience that their investors pass on through this hands-on involvement. This, in turn, increases the prospects for the success of their businesses. Indeed, entrepreneurs often report that the hands-on involvement of business angels is more valuable than the capital that they have received. However, hard evidence on the impact of this involvement on business performance remains elusive.

The informal venture capital market and the institutional venture capital market can therefore be seen as playing complementary roles in supporting entrepreneurial activity. This is evident in terms of the size and stage of investments made by business angels and venture capital funds (Freear and Wetzel, 1990). Harrison and Mason (2000) highlighted other forms of complementarity in the form of information sharing, co-investing and sequential investing and note significant collaboration in these areas between business angels and venture capital funds in the U.K. However, they also highlight the frequent tensions that arise from the different motives and expectations of angels and fund managers, the bureaucracy of venture capital funds and the unequal power relationship between angels and funds. Mason (2006) suggested that this relationship may have deteriorated during the post-2000 technology downturn.<sup>9</sup> The importance of business angels in providing a deal flow for venture capital funds is highlighted by Madill et al. (2005) who noted that 57% of technology firms in Ottawa who had received funding from angels went on to raise institutional venture capital, compared with only 10% of firms which did not raise any angel investment. It is therefore clear that a thriving institutional venture capital market requires a healthy informal venture capital market, and vice versa. Policy-makers often fail to appreciate these connections and focus their intervention on the institutional venture capital market. But as this discussion makes clear, the impact of such interventions will be compromised if the informal venture capital market is under-developed.

A third contribution of informal venture capital to economic development arises from its geographical characteristics. This has two dimensions. First, "angels live everywhere" (Gaston, 1990, p. 273). Gaston's US research suggests that the proportion of business angels in the adult population is fairly constant at around four angels in every 1000 adults. Certainly, research has documented the presence of business angels in various economically lagging regions such as Atlantic Canada (Feeney et al., 1998; Farrell, 1998; Johnstone, 2001) where institutional sources of venture capital are largely absent.<sup>10</sup> Second, various studies indicate that the majority of investments by business angels are local. This reflects both the localized nature of their business and personal networks through which they identify most of their investments (see Section 5) and their hands-on investment style and consequent need for frequent contact with their investee businesses. Two implications follow. First, in most areas outside of major financial centers and technology clusters business angels are the only source of risk capital (Gaston, 1989b). Second, the informal venture capital market is an important mechanism for retaining and recycling wealth within the region that it was created.

Informal venture capital also plays an important role in the emergence of technology clusters. This issue has attracted little explicit attention in the literature. However, it is obvious that nascent technology clusters lack

indigenous sources of institutional venture capital do not have the visibility and track record to interest venture capitalists in other cities and regions. Thus, the only source of risk capital available to technology entrepreneurs in such clusters is likely to be business angels, although, of course, they will have made their money in different (and probably mature) industries and so need to be willing to take a “punt” on businesses operating in industries that they do not understand. This was the case in the Ottawa technology cluster where the first generation of technology start-ups in the 1960s and 1970s were funded by business angels from traditional sectors (Mason et al., 2002). Once a technology cluster develops some momentum successful cashed-out technology entrepreneurs play a critical role in providing initial funding, hands-on support and credibility to the next generation of technology-based firms, grooming them for subsequent investment by venture capital funds which, by this stage in the cluster’s development are now actively investing in the cluster’s businesses. Silicon Valley, Cambridge, U.K. as well as Ottawa all provide good examples of this process.

#### 4. GOVERNMENT SUPPORT FOR THE INFORMAL VENTURE CAPITAL MARKET

This evidence on the economic significance of the informal venture capital market has prompted governments at national and state/regional scales to develop initiatives to increase investment activity by business angels. These initiatives have taken two main forms. First, evidence from early studies that business angels and entrepreneurs were incurring high search costs in trying, often unsuccessfully, to find one another on account of the fragmented nature of the market and invisibility of angels (Wetzel, 1987; Mason and Harrison, 1994), prompted the establishment of *business angel networks* (BANs). The function of these organizations—which can be thought of as being similar to “dating agencies”—is to enable entrepreneurs seeking finance to come to the attention of business angels and at the same time enable business angels to receive information on investment opportunities (filtered to meet their investment criteria if desired) without compromising their privacy (Mason and Harrison, 1996a). The pioneering BANs, such as Venture Capital Network (VCN) in New England (Wetzel and Freear, 1996) and Canada Opportunities Investment Network (COIN) in Canada (Blatt and Riding, 1996) that were established in the 1980s offered computer matching services which were intended to ensure that angels only received details of investment opportunities that matched their investment criteria. COIN started as an Ontario initiative but was extended across Canada. In the U.S., ACE-NET was created in the 1990s to enable investors to use the Internet to search for opportunities in all local/state BANs across the country

(Acs and Prowse, 2001). U.K. and continental European BANs, in contrast, have been established using investment bulletins and investment forums as their main matching mechanisms. Here again there have been attempts to forge local BANs into a national marketplace (e.g., by the U.K.'s National Business Angel Network).

BANs have received a mixed assessment. Harrison and Mason (1996) were positive about the early impact of pilot BANs in the U.K., arguing that they had mobilized capital that would otherwise have remained invisible and promoted a relatively significant number of investments which, in turn, unlocked bank lending. Entrepreneurs have also benefited from advice and signposting to more appropriate sources of assistance, feedback from investors to whom they were introduced but did not invest, while there have been wider benefits in terms of the education of entrepreneurs, investors and intermediaries and a general raising of awareness about equity. However, other evidence from the U.K. and Canada reveals mixed satisfaction with BANs among investors. Many investors report that BANs have failed to provide them with a superior quality of investment opportunities. Certainly, they have been a marginal source of investments for most angels (Blatt and Riding, 1996; Mason and Harrison, 1996b, 1999). The case for the public subsidization of BANs (Mason and Harrison, 1995) has also been challenged in the light of the willingness of private sector businesses to offer matching services. However, Mason and Harrison (1997) argued that publicly supported BANs are operating in a different part of the market than commercially oriented BANs which focus on bigger, and often later stage, investments which are able to support their fees. Meanwhile attempts in the U.K., Canada and the U.S. to create national BANs have failed on account of the strong local/regional nature of investment activity (Blatt and Riding, 1996). There is now a growing consensus that BANs need to refocus away from pure financial intermediation to a broader approach which emphasizes the education of participants in the market (Wetzel and Freear, 1996; Mason and Harrison, 1999, 2002a; Lange et al., 2002; San José et al., 2005).

Second, governments have created schemes that provide business angels with *tax incentives* in order to improve the risk-reward balance of investing in early stage businesses. Business angels are undoubtedly sensitive to levels of tax which is one of the few macro-economic factors that has a significant effect on encouraging or discouraging their investment activity (Mason and Harrison, 2000c). The U.K.'s Enterprise Investment Scheme (EIS) enables investors who make investments which qualify under the scheme's rules to write-off the amount invested against income tax. In addition, capital gains are not subject to tax, losses can be offset against tax and, perhaps most useful of all, tax that is liable on capital gains from any type of investment can be deferred if part or all of this gain is invested using the EIS. A recent evaluation of the EIS has suggested that additionality is over 50% (i.e., at least half of the monies

would not have been invested by these investors in the absence of the scheme) and that companies also benefited in terms of attracting investors who also provided business advice and expertise (Boyns et al., 2003). Several US states also offer business angels tax incentives (Lipper and Sommer, 2002). However, it is important to stress that business angels do not take the availability of tax incentives into account when evaluating *specific* investments, although this will influence how the investment is structured. For example, investments have to be in ordinary shares in order to qualify for EIS relief, even though current best practice suggests that preference shares may be a more appropriate investment instrument.

These initiatives have been supplemented by *amendments to securities legislation* which control the promotion of share issues in order to provide investor protection. Firms wishing to raise finance from the general public are required to produce a prospectus which has been approved by an authorized organization to ensure that they are not potentially misleading. However, the costs involved are too high for the typical fund-raising exercise. This is no longer necessary in several countries, such as the U.S., Australia and, most recently, the U.K., if the offer is promoted to self-certified high net worth individuals or sophisticated investors who give up certain legal protections and channels of legal redress to receive investment opportunities (HM Treasury, 2004a, b). BANs in the U.K. have been exempt from the regulations concerning the promotion of investments for some time (Clarke, 1996).

However, in the light of recent evidence that business angels continue to be opportunity constrained despite being members of BANs (Mason and Harrison, 2002a), it is now recognized that there are also demand-side barriers to investment. A lot of businesses looking for investment from business angels are not investment ready, with missing information in the business plan (e.g., competitor analysis) and poorly developed ideas about the business model, markets, route to market and unrealistic expectations about investor requirements (e.g., involvement) (Feeney et al., 1999; Mason and Harrison, 2001, 2004a). These deficiencies are often accompanied by poor presentation (Mason and Harrison, 2003). Accordingly, recent interventions have sought to address the issue of “investment readiness” (Mason and Harrison, 2001). There are examples of investment ready programs in Canada (Industry Canada, 2001) and the U.K. (SQW, 2004). An alternative approach of *investment facilitation* is discussed by Mason and Harrison (2004a).

The most recent form of initiative is *co-investment schemes*. This has been prompted by the post-2000 venture capital investment downturn which followed the collapse of the technology bubble of the late 1990s. The response of venture capital firms was to cut back on making new investments in order to focus their attention on the businesses in their existing portfolios. The consequence for business angels was that they were unable to pass on those businesses

in their portfolios to venture capital funds for follow-on investments and so had to do more follow-on investments themselves. This meant that they had less money and time available to make new investments. Co-investment schemes have addressed this liquidity constraint by matching angel investments with public money on a one-to-one basis up to a maximum figure. Angels have also co-invested alongside technology programs such as SBIR and the Advanced Technology Program in the U.S. (Chang et al., 2002) and SMART in the U.K. which provide grants to technology companies to make the transition from the laboratory to the market place. These schemes are particularly attractive to business angels. First, the funds provide a means of risk sharing. Second, the competitive peer review process by technology and business experts provides an independent source of assessment which assists in the due diligence process (Sohl, 2003).

## 5. THE INVESTMENT PROCESS

The aim of the early studies of the informal venture capital market was, in the words of William Wetzel Junior, the pioneer of the field, “to put boundaries on our ignorance” (Wetzel, 1986, p. 132) by generating insights into the characteristics of business angels and their investment activity. In contrast, ‘second generation studies’ have focused on the investment process (Mason and Harrison, 2000a). Following Riding et al. (1993) and Haines et al. (2003) a number of discrete stages can be identified (see Table 10-3):

- Deal origination.
- Deal evaluation: this can, in turn, be sub-divided at least two sub-stages:
  - initial screening,
  - detailed investigation.
- Negotiation and contracting.
- Post-investment involvement.
- Harvesting.

This sequence is similar in most respects to the investment decision-making model of institutional venture capital funds (Tyejee and Bruno, 1984; Fried and Hisrich, 1994). However, the approach of business angels is less sophisticated.

Agency theory provides a framework to study the investment process. An agency relationship is said to exist when one individual (the principal) engages the services of another individual (the agent) to perform a service on their behalf (Jensen and Meckling, 1976). This involves the delegation of a measure of decision-making authority from the principal to the agent. Both



TABLE 10-3 *Stages in the business angel's investment decision*

|                             |   |
|-----------------------------|---|
| Deal origination            | The investor becomes aware of the opportunity—typically through one of the following channels: chance encounter, referral from business associates or other individuals or organizations in their network, or personal search   |
| Deal evaluation             | Two stages:<br>(i) Initial screening/first impressions: key considerations are the 'fit' with the investor's personal investment criteria, their knowledge of the industry/market and their overall impression of the potential of the proposal. Also influenced by the source of the referral<br>(ii) Detailed evaluation: the investor will examine the business plan in detail, consult with associates, will meet the principals, take up references, research the proposal. The decision will be influenced by the potential of the industry, the business idea, impressions of the principals and potential financial rewards |
| Negotiation and contracting | Negotiations with the entrepreneur over valuation, deal structuring and the terms and conditions of the investment. Main factor is pricing  |
| Post-investment involvement | Investor is likely to become involved with the business in some kind of hands-on capacity, including advice and mentoring, networking, functional input and member of board. Degree of involvement may vary according to the stage of business development and the performance of the business  |
| Harvesting                  | Exit from the business, either because it fails or by selling their shares to another investor. Investors normally exit from successful investments by means of a trade sale  |

are assumed to be economic-maximizing individuals. The central concern of agency theory is opportunism. The separation of ownership and control creates the risk that the agent will make decisions that are not in the best interests of the principal. This creates two types of risk for the principal (i.e., the investor). The first is adverse selection which arises as a result of informational asymmetries: the agent is better informed than the principal about their true level of ability. However, agents may deliberately misrepresent their abilities to the principal. The second risk is moral hazard. In situations where it is not possible for the principal to observe the behaviors of agents the agent may shirk, engage in opportunistic behavior that is not in the interests of the principal or pursue divergent interests that maximize their economic interests rather than those of the principal. Fiet (1995) argued that every investment decision also includes market risk—the risk that the business will perform less well than anticipated on account of competitive conditions (e.g., competition, demand, technological change). This section considers how business angels manage these sources of risk.

### 5.1. Deal Origination

The evidence is consistent in suggesting that business angels adopt a relatively *ad hoc* and unscientific approach to identifying investment opportunities. Atkin and Esiri (1993) emphasized that most investments arise from chance encounters. Informal personal contacts—business associates and friends—are the most significant sources of deal flow. Professional contacts are much less significant: of these, accountants are the most frequent sources whereas few business angels receive deal flow from lawyers, bankers and stockbrokers. Those angels who are known in their communities also receive approaches from entrepreneurs. Information in the media is another source of deal flow for a significant minority of business angels. Some business angels also undertake their own searches for investment opportunities. Those business angels who are members of BANs also report that they are significant sources of deal flow (Mason and Harrison, 1994, 2002a). In some cases—especially in the case of *ad hoc* investors—the entrepreneur is not a stranger but a business associate who is known to the angel (e.g., client, supplier) (Atkin and Esiri, 1993). Kelly and Hay (2000) observed that the most active investors have less reliance than occasional investors on “public” sources (e.g., accountants, lawyers, etc.) for their deal flow and place more emphasis on “private” sources. Thus, most of their deals are referred by individual and institutional sources in their extensive and longstanding networks of relationships.

However, these various sources of information differ in their effectiveness. Freear et al. (1994b) calculated yield rates for various sources of deal flow (i.e., comparing investments made against deals referred for each information source). This points to the informal personal sources of information—business associates, friends and approaches from entrepreneurs—as the ones that have the highest probability of leading to investments whereas non-personal sources such as accountants, lawyers and banks have a low likelihood of generating investments. These findings are largely corroborated by Mason and Harrison (1994) for the U.K. However, in their study the highest yield rates are recorded by some of the infrequently used professional contacts, notably banks and stockbrokers. This study also notes the low yield ratio for BANs. Riding et al. (1995) found that the rejection rate at the initial screening stage for deals referred by business associates is lower than that for other referral sources.

Investing in businesses that are referred by trusted business associates and friends is an obvious way in which business angels can minimize adverse selection problems. As Riding et al. (1995) commented, “even if the principals of the firm are unknown to the investors, if the investor knows and trusts the referral source risk is reduced.” Deal referrers are passing judgment on the merits of the opportunity and so are putting their own credibility and reputation on the line.

## 5.2. Deal Evaluation

The process of evaluating investment opportunities involves at least two distinct stages—initial screening and detailed investigation (or due diligence—Riding et al., 1993)—although this is not reflected in most studies. The initial step of business angels is to assess investment opportunities for their “fit” with their own personal investment criteria. The investment opportunity will also be considered in terms of its location (how close to home?), the nature of the business and the amount needed and any other personal investment criteria (Mason and Rogers, 1997). The business angel will also typically ask themselves two further critical questions: first, “Do I know anything about this industry, market or technology?” and second, “Can I add any value to this business?” Clearly, the ability to add value is very often a function of whether the angel is familiar with the industry. If the answer to either question is negative then the opportunity will be rejected at this point.

Angels then undertake a quick review of those opportunities that fall within their investment criteria to derive some initial impressions. Although most business angels expect a business plan, they are unlikely to read it in detail at this stage. Their aim at this point in the decision-making process is simply to assess whether the proposal has sufficient merit to justify the investment of time to undertake a detailed assessment. This stage has been the subject of a detailed analysis by Mason and Rogers (1996, 1997) using verbal protocol analysis, an experimental-type technique which asks subjects (in this case business angels) to think out loud as they perform a task (in this case evaluating a real investment opportunity). They observe that angels approach this stage with a negative mindset, expecting that the opportunity will be poor (because of the opportunities that they have previously seen) and looking for reasons to reject it. This approach has been termed “three strikes and you’re out” (Mason and Rogers, 1996, 1997) and is supported by evidence that the rejection of opportunities is generally based on several factors rather than a single deal killer (Mason and Harrison, 1996c). The market and the entrepreneur are the key considerations at this stage. Less significant are the product/service and financial factors. Indeed, angels exhibit considerable skepticism about the value of financial information in the business plan of start-ups: as one investor in the Mason and Rogers (1996, p. 45) study commented, “I take [financial projections] with a great pinch of salt, especially from accountants because they can tweak the assumptions and come up with any figure. So, it’s the last thing I look at.” Nevertheless, investors want to see that there is the potential for significant financial return, that the principals are financially committed and what the money that is invested will be used for. Some angels will be flexible, willing to treat these criteria as compensatory (e.g., a strong management team would compensate for a distant location), whereas others will regard them as noncompensatory (Feeney et al., 1999).

The purpose of the initial screen is to filter out “no hopers” in order to focus their time on those opportunities that appear to have potential. These are subject to more detailed appraisal. The investor will read the business plan in detail, go over the financial information, visit the premises, do some personal research to gather additional information on market potential, competition and so on, and assess the principals. Indeed, getting to know the principals personally (by a series of formal and informal meetings) is the most vital part of the process (May and Simmons, 2001). This stage has received little attention from researchers. According to May and Simmons (2001, p. 101), “it might consist of a few phone calls and a visit or two, or weeks of meetings, documents flying back and forth and questions, questions, questions.” However, it would appear that most angels emphasize their intuition and gut feeling rather than performing formal analysis (Haines et al., 2003)—although more experienced angels, and angel groups (see Section 6) adopt more sophisticated approaches (e.g., see Blair, 1996).<sup>11</sup>

Once the opportunity has passed from the initial screen the importance of “people” factors becomes critical (Riding et al., 1995), with investors emphasizing management abilities, an understanding of what is required to be successful, a strong work ethic, integrity, honesty, openness and personal chemistry (Haines et al., 2003; Mason and Stark, 2004). This reflects the long and personal nature of the angel-entrepreneur relationship.<sup>12</sup> Rewards, realism of the projections and potential also assume greater importance while “investor fit” becomes less of a consideration (Riding et al., 1995).

This stage ends when the investor has decided whether or not to negotiate a deal with the investor. In their Canadian study Riding et al. (1993) found that 72.6% of opportunities were rejected at the initial impressions stage, a further 15.9% were rejected following more detailed evaluation, and as this stage proceeds another 6.3% were eliminated, a cumulative rejection rate of 94.8%. Thus, business angels proceed to the negotiation stage with only 5% of the investment opportunities that they receive.

The key role of the entrepreneur/management team in the decision whether or not to invest is confirmed in other studies. Using conjoint analysis—a method to measure quantitatively the relative importance of one decision-making criteria in relation to another (see Shepherd and Zacharakis, 1999)—Landström (1998) found that business angels attach the greatest importance to the leadership capabilities of the principals, followed by the potential of the firm’s market and products. Feeney et al.’s (1999) approach was to ask business angels “what are the most common shortcomings of business opportunities that you have reviewed recently?” This highlighted shortcomings in both the management (lack of management knowledge, lack of realistic expectations, personal qualities) and the business (poor management team, poor profit potential for the level of risk, poor fit, undercapitalized/lack of liquidity,

insufficient information provided). Asking investors “what are the essential factors that prompted you to invest in the firms that you have chosen?” (Feeney et al., 1999) highlighted three management attributes—track record, realism and integrity and openness—and four attributes of the business—potential for high profit, an exit plan, security on their investment and involvement of the investor. However, while the primary deal killer is the perception of poor management, the decision to invest in an opportunity involves a consideration of management ability, growth and profit potential. In other words, angels are looking for businesses that show growth potential and have an entrepreneurial team with the capability to realize that potential (Feeney et al., 1999). Both these studies also emphasize that investment criteria are personal, with angels using different criteria in their assessment of investment proposals. For example, Feeney et al. (1999) suggested that the decision processes of more experienced investors differs from that of less experienced investors.

This emphasis on the entrepreneur reflects the view of angels that agency risk is more of a threat than market risk. Fiet (1995) argued that business angels lack information or the tools and resources to evaluate market risks effectively. As a consequence, they specialize in evaluating agency risk—assessing whether or not the entrepreneur can be relied upon as a venture manager—while relying upon competent and trustworthy entrepreneurs to manage market risk.<sup>13</sup> This contrasts with venture capital funds which attach more importance to market risk than agency risk. They are less concerned with agency risk because they have learned how to protect themselves using stringent boilerplate contractual provisions which allows them to replace an entrepreneur who is not performing or is found to be incompetent. Thus, “compared with venture capital investors, business angels place much more importance upon screening entrepreneurs than deals for market risk” (Fiet, 1995, p. 567).

### 5.3. *Negotiation and Contracting*

Having decided, in principle, to invest the business angel must negotiate terms and conditions of the investment that are acceptable both to themselves and also to the entrepreneur. There are three main issues—valuation, structuring of the deal (share price, type of shares, size of shareholding, timing) and the terms and conditions of the investment, including the investor’s role. In agency theory terms deal structuring—mechanisms for allocating the rewards to the investor and entrepreneur—are an attempt to align the behaviors of the entrepreneur with that of the investor, while the terms and conditions attempt to control the behaviors of the entrepreneur. These are major lacunae in the informal venture capital market research.

In the study by Riding et al. (1993), half of the investment opportunities that reached this stage were not consummated. The most frequent reason for not

making an investment was associated with valuation, notably “inappropriate views by entrepreneurs (in the opinion of the investors) regarding the value of the firm as a whole and, within the firm, the value of an idea compared to the overall value of a business. Most investors note that potential entrepreneurs overvalue the idea and undervalue the potential contributions (both financial and non-financial) that are required to grow and develop a business” (Haines et al., 2003, p. 24). Putting a value on the “sweat equity” of the entrepreneurs is also problematic.

There is no universally agreed method of valuing a small company. Market-based valuations are inappropriate because small businesses are not continually valued by the market and appropriate comparator stocks are unlikely to be available. Asset-based valuations are more commonly used although finance theory prefers earnings or cash-flow based valuations because they value the business in terms of the future stream of earnings that shareholders might expect from the business. However, these approaches are complex. Valuation of new and early stage businesses adds further complications because they may only have intangible assets (e.g. intellectual property). It is therefore not surprising, especially since most angel investments are concentrated at start-up and early stage, that methods of pricing and calculating the size of shareholdings are remarkably imprecise and subjective (Mason and Harrison, 1996d), based on rough rules of thumb or gut feeling. As investors, May and Simmons (2001, p. 129) noted that “the truth about valuing a start-up is that it’s often a guess.” Where an attempt is made to price the investment on a more rigorous basis then the earnings based approach is the most common method (Lengyel and Gulliford, 1997).

Angels draw up contracts as a matter of course to safeguard their investment, although their degree of sophistication varies. Contracts specify the rights and obligations of both parties and what will be done, by whom and over what time frame. Their objective is to align the incentives of the entrepreneur and the investor by means of performance incentives and direct control measures. Kelly and Hay (2003) noted that certain issues are nonnegotiable: veto rights over acquisitions/divestments, prior approval for strategic plans and budgets, restrictions on the ability of management to issue share options, noncompete contracts required by entrepreneurs on the termination of their employment in the business and restrictions on the ability to raise additional debt or equity finance. These issues give investors a say in material decisions that could impact the nature of the business or the level of equity holding. However, there are also a number of contractual provisions to which angels attach low importance, and which might be considered to be negotiable. These include forced exit provisions, investor approval for senior personnel hiring/firing decisions, the need for investors to countersign bank checks, management equity ratchet provisions and the specification of a dispute resolution mechanism. Less experienced

investors place relatively greater emphasis on the need to include a broad array of contractual safeguards to protect their interests. However, experienced investors are more likely to include specific provisions that can impact the level of their equity stake (share options, ratchets) and the timing of exit (forced exit provisions). In other words, with experience business angels become more focused on those elements that can impact their financial return.

Investors recognize that the investment agreement must be fair to both sides (May and Simmons, 2001): contracts that favor the investor will be detrimental to the entrepreneur's motivation. In Mason and Harrison's (1996d) study, two-thirds of investors and entrepreneurs considered that the investment agreement was equally favorable to both sides, and half of the investors reported that this was their objective. Indeed, a significant minority of investors believed that the agreement actually favored the entrepreneur. Thus, the available evidence suggests that in most cases entrepreneurs are not exploited by investors when raising finance.

The inclusion of contractual safeguards does not indicate whether investors will be willing to invoke them to protect their interests. Moreover, contracts are, of necessity, incomplete by their very nature. There are three reasons for this: it is costly to write complete contracts; it is impossible to foresee all contingencies; and on account of asymmetric information (van Osnabrugge, 2000). Thus, in practice investors place a heavy reliance on their relationship with the entrepreneur to deal with any problems that arise (van Osnabrugge, 2000; Kelly and Hay, 2003). Indeed, Landström et al. (1998) argued that one of the purposes of establishing a contractual framework at the outset is to provide a basis for the development of a relationship between the parties to develop. In other words, the contract is less a protection mechanism *per se*; rather, it is a means by which mutual behaviors expectations of all parties in the transaction can be clarified.

Most angel investments involve input from professional advisers. For example, lawyers would normally review, and might draw up, the investment agreement, but would not be involved in the negotiations. Similarly, accountants may be consulted for advice but would rarely play a more prominent role. Thus, transactions costs are low (Mason and Harrison, 1996d). In Lengyel and Gulliford's (1997) study the entrepreneur's costs amounted to an average of 5.1% of the funds raised (and 29% reported no costs) while for the investor the average costs were 2.8% of the amount invested (and 57% reported no costs).

The time taken by business angels to make investments is much quicker than that of venture capital funds (Freear et al., 1995). Mason and Harrison (1996d) reported that in their study the entire investment process rarely extended over more than three months, and often took less than a month. Most negotiations took less than a week to complete whereas the evaluation could take up to three months or more. Thus, in nearly half of the investments less

than a month elapsed between the entrepreneur's first meeting with the investor and the decision to invest; in 85% of cases the elapsed time was under three months.

#### *5.4. Post-Investment Involvement*

From an agency perspective, monitoring is the main way in which principals attempt to mitigate the risk of opportunistic behaviors on the part of the agent going undetected. In line with this expectation, most business angels play an active role in their investee businesses. There is a spectrum of involvement: at one extreme are passive investors who are content to receive occasional information to monitor the performance of their investment while at the other extreme are investors who use their investment to buy themselves a job. However, most angels do not want day-to-day involvement hence the typical involvement ranges from a day a week (or its equivalent) to less than a day a month (Mason and Harrison, 1996d). Nevertheless, Sætre (2003) emphasized that some angels are so involved, and involved so early, that they are indistinguishable from the entrepreneurs, and are seen by the entrepreneurs as being part of the entrepreneurial team. In a similar vein, Politis and Landström (2002) see angel investing as simply a continuation of an entrepreneurial career.

Madill et al. (2005) identified a number of roles that business angels play in their investee businesses: advice about the management of the business, contacts, hands-on assistance (e.g., legal advice, accountancy advice, provision of resources), providing business and marketing intelligence, serving on the Board of Directors or Advisory Board, preparing firms to raise venture capital and providing credibility and validation. Sørheim (2005) emphasized the role of business angels in helping their investee businesses to raise additional finance. The nature and level of involvement is influenced by geography. Landström (1992) noted that frequency of contact between angels and their investee companies is inversely related to the geographical distance that separates them. It will also be influenced by the performance of the business, with angels more involved at particular stages of business development and in crisis situations.

However, in contrast to agency theory the involvement of angels in their investee businesses is not motivated by monitoring considerations. First, as noted earlier, angels derive psychic income from their involvement in their investee businesses in the form of fun and satisfaction from working with new and growing businesses and their belief that their experience, know how and insights can "make a difference." May and Simmons (2001, p. 156) quoted one investor as follows: "I've never had as much fun in my life. It's a joy to see someone listen, take action and win." Second, angels see themselves as "offering help" rather than "checking up" on their investee businesses by acting as mentors, providing contacts, guidance and hands-on assistance (Haines et al.,



2003). Third, as Kelly and Hay (2003, p. 309) commented, “from the outset, the relationship between the business angel and the entrepreneur appears to be more positive and trusting in character than the inherently adversarial one implied by agency theorists.”

A majority of entrepreneurs and angels regard their relationship as productive and consensual—although entrepreneurs have a more favorable view of its productiveness than angels (Freear et al., 1995; Mason and Harrison, 1996d). One study reported that half of the entrepreneurs who had raised finance from business angels regarded their contributions as being helpful or very helpful (Mason and Harrison, 1996d). Another study reported that entrepreneurs considered that the most valuable contribution of their business angel has been as a sounding board (Harrison and Mason, 1992). There is a suggestion that entrepreneurs want their investors to be more involved in certain areas, especially financial management (Ehrlich et al., 1994). Criticisms by entrepreneurs who have raised finance from angels are mainly concerned with those who lack knowledge of the product or market (Lengyel and Gulliford, 1997). Finally, most business angels report that they have derived fun and enjoyment from their investments, often more than expected, in cases where the investment is still trading, but not when the business has failed. Psychic income returns are therefore related to business performance rather than compensating for financial loss (Mason and Harrison, 1996d). However, there has been no rigorous attempt to assess whether this involvement of business angels has a favorable impact on the performance of their investee businesses. There is no evidence from research on venture capital funds that greater involvement is a necessary condition for adding value nor whether involvement produces enhanced business performance (Sapienza and Gupta, 1994; Fried et al., 1998). This may be because the involvement of venture capitalists is concentrated on their poorly performing investments, determining whether and how they can be turned around, or even whether continued support is desirable (Zider, 1998; Higashide and Birley, 2002).<sup>14</sup> There are also formidable methodological challenges.<sup>15</sup>

### 5.5. *Harvesting*

Investing in unquoted companies is regarded as high risk. Certainly, the performance of European venture capital funds specializing in early stage investments (in practice this means technology-focused investments) have much lower rates of return than those which focus on later stage investments (EVCA, 2005). Diversification is the main strategy for reducing risk. However, this is not an option for business angels. First, typically they have just a handful of investments in their portfolios. Second, they often restrict their investments to sectors which they know and understand, so their portfolios are unbalanced.

Third, as the first external investor in a business, and generally lacking the financial resources to make follow-on investments, they are vulnerable to being diluted in the event that further funding rounds are required.

There have been only two studies of the investment returns of business angels, a small scale Finnish study (Lumme et al., 1998) and a larger U.K. study (Mason and Harrison, 2002b). It is important to note that these studies only measure multiples achieved on the amounts invested. However, many angels also attempt to draw back at least part of their investment in the form of a director's fee or interest on loans provided, either immediately or at some stage in the future when the business is financially stronger. This could be quite a significant proportion of the investment in smaller deals (Mason and Harrison, 1996d; Lengyel and Gulliford, 1997). The U.K. study highlights the highly skewed distribution of returns, with 40% of investments making a loss (34% a total loss), and another 13% only achieving break-even or generating bank account-level returns. However, there was a significant subset of investments, some 23% in total, which generated internal rates of return (IRRs) in excess of 50%.

The U.K. study went on to explore the types of *investments* that were likely to be successful. It identified large investments, large deal sizes and deals involving multiple investors as being more likely to be high performing investments (Mason and Harrison, 2002b). A separate analysis of the returns distribution of technology and nontechnology investments found no significant differences in the returns profile (Mason and Harrison, 2004b). This may suggest that the risk of investing in technology sectors has been over-stated. Another possibility is that business angels are better able to mitigate the risks involved in investing in technology businesses on account of their specialist expertise and entrepreneurial background.

The Finnish study, in contrast, sought to identify differences between the most, and least, successful *investors*. The most successful investors were more likely to be motivated by the fun and interest of making such investments, have a large deal flow and have a lower estimation of the value of their hands-on involvement. The least successful investors were more likely to be motivated by altruism, have a low deal flow and make few investments and rely to a greater extent on friends for deal flow. They were also more likely to make investments in friends' businesses and have a different pattern of hands-on involvement, over-emphasizing contributions that other research has suggested are least important in adding value (Lumme et al., 1998).

Comparison with the returns achieved by institutional venture capital investors is problematic because the reporting unit is the fund, whereas angels invest on a deal-by-deal basis. However, Murray (1999) reported deal-specific returns for one U.K. venture capital fund. Comparing the returns achieved by business angels with this information reveals a much higher loss rate by the

venture capital fund (64%) and a lower proportion of investments that generated a moderate return, but a very similar proportion of high return investments (IRR in excess of 50%). The interpretation of these differences is that because the venture capital fund is seeking to maximize the performance of the fund it can be more ruthless with those investments that are performing moderately, in order to focus the time of its executives on supporting the best performing investments whereas business angels invest on a deal-by-deal basis (Mason and Harrison, 2002b).

Business angels are thought to be relatively patient investors, willing to hold their investments for up to seven years or more (Wetzel, 1981; Mason and Harrison, 1994). In reality, angels hold their investments for a much shorter time. The median time to exit in the U.K. is four years for high performing investments and six years for moderately performing investments, while failures appear, on average, after two years (Mason and Harrison, 2002b). In Finland investments that had a positive outcome were five years old at harvest whereas those that failed had an average holding time of 2.8 years. In both studies a trade sale (i.e., sale of the company to another company) was the most common exit route for successful investments, with an IPO only accounting for a small minority of cases. Trade sales, along with sale to existing shareholders were the most common exit routes for investments with little or no value.

## 6. THE EVOLUTION OF THE ANGEL MARKET

Recent research in the U.S. has revealed that the angel market place is evolving from a largely invisible, atomistic market dominated by individual and small *ad hoc* groups of investors who strive to keep a low profile and rely on word-of-mouth for their investment opportunities, to a more organized market place in which angel syndicates (sometimes termed ‘structured angel groups’) are becoming increasingly significant. As a result, the angel market place is in the process of being transformed from a “hobby” activity to one that is now increasingly professional in its operation, with published routines for accessing deals, screening deals, undertaking due diligence, negotiating and investing (May, 2002). Sohl et al. (2000) claimed that “angel alliances are the fastest growing segment of the early stage equity market.” However, solo investors still dominate the market (Lengyel and Gulliford, 1997; Investor Pulse, 2003; Infometrics, 2004).

The Band of Angels, which was founded in Silicon Valley in 1995, is generally regarded as the first organized syndicate to be formed. Others, such as Tech Coast Angels (1997), Sierra Angels (1997), Common Angels (1997) and The Dinner Club (1999), soon followed.<sup>16</sup> There are currently estimated to be around 200 angel syndicates located throughout the U.S. and growing evidence

of specialization by industry sector (e.g., health care angel syndicates) and type of investor (e.g., women-only angel syndicates). A national body to bring angel groups together for the purposes of transferring best practice, lobbying and data collection was created in 2003 (Angel Capital Association, 2005). The same trend is also clearly evident in the U.K. although at an earlier stage, and it has not attracted the same degree of attention from researchers or commentators.<sup>17</sup>

Angel syndicates emerged because individual angels found advantages of working together, notably in terms of better deal flow, superior evaluation and due diligence of investment opportunities, and the ability to make more and bigger investments, as well as social attractions. They operate by aggregating the investment capacity of individual high net worth individuals (HNWIs). Some groups are member-managed while others are manager-led (Preston, 2004). Syndicates take various forms but the most common generic type of model (at least in the U.S.) is as follows:

- Limited and selective membership of angels (typically 20–75 members) who typically play an active role in the investment process.
- Meet regularly (e.g., for dinner) to hear ‘pitches’ by entrepreneurs seeking finance.
- A syndicate manager supports members by organizing meetings, communications and manages logistics.
- The manager or a core group of members will screen the deal flow and select the companies which are invited to pitch.
- Q&A session follows each pitch.
- Angels vote whether to pursue their interest in the business.
- If the vote is in favor a sub-group will be appointed to undertake the due diligence and report back to the full membership.
- If the recommendation is positive, individual members make their own decisions whether or not to invest (there is likely to be a minimum investment threshold for each deal) and the syndicate will combine all of the member dollars into a single investment. Alternatively, if the syndicate operates a pooled fund a majority vote will decide whether or not to invest.
- An expectation that each member of the syndicate will make a certain number of investments per year.

Some of the larger and longer established U.S. syndicates have also established sidecar funds—that is, committed sources of capital that invest alongside the angel group. The investors in such funds are normally the syndicate members but may also include other HNWIs or institutions. These funds give the syndicate additional capital to invest in deals to avoid dilution, enables syndicate members to achieve greater diversification by exposing them

to more investments than they can make directly through the syndicate, and is a means of attracting “right-minded” investors who want to participate in seed and early stage deals but cannot be active members of a syndicate (e.g., because of lack of time).

The emergence of angel syndicates is of enormous significance for the development and maintenance of an entrepreneurial economy. First, they reduce sources of inefficiency in the angel market. The angel market has traditionally been characterized by inefficiency on account of the fragmented and invisible nature of angels. There was no mechanism for angels to receive a steady flow of investment opportunities. They found their deals by chance. The entrepreneur’s search for angel finance was equally a hit-or-miss affair. Investors and entrepreneurs both incurred high search costs (Wetzel, 1987; Mason and Harrison, 1994). This encouraged many to drop out of the market as either suppliers or seekers of finance. Angel syndicates, in contrast, are generally visible and are therefore easier for entrepreneurs to approach.

A further source of inefficiency was that each investment made by an investor has typically been a one-off that was screened, evaluated and negotiated separately. However, because of the volume of investments that angel syndicates make they have been able to develop efficient routines for handling investment inquiries, screening opportunities and making investment agreements.

Second, they have stimulated the supply-side of the market. Syndicates offer considerable attractions for HNWIs who want to invest in emerging companies, particularly those who lack the time, referral sources, investment skills or the ability to add value. However, many individuals who have the networks and skills to be able to invest on their own are also attracted by the reduction in risk that arises from investing as part of a syndicate, notably the ability to spread their investments more widely and thereby achieve greater diversification, and access to group skills and knowledge to evaluate investment opportunities and provide more effective post-investment support. Other attractions of syndicates are that they enable individual angels to invest in particular opportunities that they could never have invested in as individuals, offer the opportunity to learn from more experienced investors and provide opportunities for camaraderie and schmoozing with like-minded individuals. Syndicates will also be attractive to individuals who want to be full-time angels. Thus, angel syndicates are able to attract and mobilize funds that might otherwise have been invested elsewhere (e.g., property, stock market, collecting), thereby increasing the supply of early stage venture capital, and to invest it more efficiently and effectively.

Third, they are helping to fill the “new” equity gap. Venture capital funds have consistently raised their minimum size of investment and are increasingly abandoning the early stage market (after briefly returning during the “dot-com bubble” of the late 1990s). Most funds have a minimum investment size of at least £500,000 and the average early stage investment by U.K. venture

capital funds in recent years has been around £1m (BVCA, 2004). This has resulted in the emergence of a new equity gap—roughly the £250,000 to £2m+ range which covers amounts that are too large for typical “3F” money (founder, family, friends) but too small for most venture capital funds. Angel syndicates are now increasingly the only source of venture capital in this size range. The same trends—increasing deal sizes by venture capital funds and emergence of angel syndicates to fill the gap—are also evident in the U.S. where the “gap” is estimated to be in the \$500,000 to \$5m range (Sohl, 1999, 2003).

Fourth, they have the ability to provide follow-on funding. One of the potential problems of raising money from individual business angels is that they often lack the financial capacity to provide follow-on funding. The consequence has been that the entrepreneur is often forced to embark on a further search for finance. Moreover, in the event that the need for additional finance is urgent then both the entrepreneur and the angel will find themselves in a weak negotiating position with potential new investors, resulting in a dilution in their investments and the imposition of harsh terms and conditions. With the withdrawal of many venture capital funds from the small end of the market individual angels and their investee businesses have increasingly been faced with the problem of the absence of follow-on investors. However, because angel syndicates have got greater financial firepower than individual angels or *ad hoc* angel groups they are able to provide follow-on financing, making it more efficient for the entrepreneur who avoids the need to start the search for finance anew each time a new round of funding is required.

Fifth, their ability to add value to their investments is much greater. The range of business expertise that is found among angel syndicate members means that in most circumstances they are able to contribute much greater value-added to investee businesses than an individual business angel, or even most early stage venture capital funds. May and Simmons (2001, p. 156), leading angel syndicate practitioners in the U.S., commented that “when angels band together . . . their smorgasbord of advice and strategic services frequently makes the difference between life and death for a start-up.”

Finally, angel syndicates have greater credibility with venture capitalists. Venture capital funds often have a negative view of business angels, seeing them as amateurs whose involvement in the first funding round of an investment could complicate subsequent funding rounds because of their tendency to over-price investments, use complicated types of investment instruments and make over-elaborate investment agreements (Harrison and Mason, 2000). Venture capitalists may therefore avoid deals in which angels are involved because they perceive them to be too complicated to do. However, because of the professionalism and quality of the membership of angel syndicates venture capital funds hold them in much higher esteem. Accordingly, the increasing prominence of angel syndicates results in much greater complementarity between the angel

market and venture capital funds, to the benefit of fast-growing companies that raised their initial funding from angel syndicates but now need access to the amounts of finance that venture capital funds can provide.

## 7. CONCLUSION

This chapter has sought to highlight the significance of the informal venture capital market as a source of funding for entrepreneurial businesses. However, its significance is frequently overlooked in both the academic and practitioner literature and by policy-makers where the emphasis continues to be placed on institutional venture capital, despite its almost non-existent role in funding new and recently started businesses. There are three inter-related reasons why the informal venture capital market is often overlooked. First, the market is invisible and fragmented. There are no directories of angel investors and their investments are not recorded in any systematic way. Second, because of the invisibility of business angels, and their efforts to maintain their secrecy, it is extremely difficult to undertake research on the size and operation of the market. Research is typically based on small-scale snapshot samples of convenience which are unsuited to statistical analysis. Third, the research base is limited and largely atheoretical. Indeed, the initial studies in the 1980s and early 1990s were descriptive, aimed at profiling angel characteristics, motivations and investment activity. However, recent research has become more analytical, focusing on actual behaviors rather than preferences, on aspects of the investment process rather than on the actors, and has become more anchored in theory, with several studies using agency theory as a framework for analysis. Nevertheless, the opportunities for further research are considerable.

First, there is an urgent need to get away from snapshot surveys of the angel market and to develop longitudinal research on the angel market. This involves two dimensions (Sohl, 2003). The first has the business angel as the unit of analysis and seeks to develop information on investor and investment trends. The challenge, as always, is in the methodology. One approach is simply to repeat snapshot surveys at regular intervals. A more manageable, if partial, approach is to identify and survey angel syndicates on a regular basis, while a third approach would be to develop an angel panel which is surveyed on a regular basis. The second approach takes the deal as the unit of analysis and tracks it from the point of referral to the angel through to rejection or investment and on to subsequent funding round and exit. Much of the research in venture capital is 'timeless' in the sense that it does not reflect the economic conditions of the time (Mason and Harrison, 2004c). Thus, an important dimension of such longitudinal studies involves relating investment trends to the wider economic conditions of the time. For example, how angels responded to the post-2000

investment downturn remains largely unexplored and unanswered (but see Mason, 2006, for a brief discussion).

Second, the emergence of angel syndicates raises a series of questions. Are they attracting investors who are new to the market, and thus new money that would otherwise have been invested elsewhere, or are they attracting solo angels? If they are attracting solo angels, will this deplete the population of small-scale investors and thereby re-open the sub £250,000/\$500,000 equity gap? As angel syndicates become more organized and develop fixed costs will this lead to an upward drift in their investment activity, thereby re-opening the equity gap?

Third, taking a “food chain” perspective, are the complementarities between angels and venture capital funds diminishing as venture capital funds continue to shift their investment focus to larger and later stage deals. Can angel syndicates fill this gap—are their financial resources big enough to by-pass venture capital funds and take their investee businesses to a harvest event themselves, or by co-investing with other angel syndicates? Indeed, are we seeing the beginning of a bifurcation of the venture capital market between businesses that because of the scale of their R&D or capital investment require multi-million dollar investments over several rounds (e.g., life sciences, telecoms infrastructure), and therefore need funding from venture capital funds, and businesses (e.g., software) whose funding requirements are more modest, in the \$10m–\$20m range, and so could be funded largely or entirely by angel syndicates?

Fourth, the chapter has noted that many governments now recognize the economic significance of business angels and have introduced various measures to support the informal venture capital market. However, Aernoudt (1999) argued that the case for government intervention is not proven. Thus, there is scope for further applied research which explores whether the case for intervention is justified, and if the case is supported what is the most appropriate form(s) of intervention. Can the studies that various national venture capital associations undertake of the economic impact of venture capital be replicated for informal venture capital? Research from various countries is consistent in finding that angels are opportunity constrained. Understanding the reasons would seem to be the top priority for policy-makers. How much stems from the limitations of the investors themselves (e.g., restricted investment criteria, competence limitations), how much is due to the inefficiencies in the operation of the market and how much is a result of the lack of investment readiness among businesses seeking finance? Can “second generation” business angel networks—which focus on raising the competence of the participants in the market—make a difference?

There are also a host of issues where information is either lacking or requires corroboration. Examples of the former include identifying the



characteristics of altruistic investors (Sullivan, 1994), women business angels (Brush et al., 2002; Harrison and Mason, 2005) and successful investors. The concept of an “angel career” (Politis and Landström, 2002) offers a potentially useful way in which to explore angel learning. The negotiation, valuation and contracting stages remain poorly understood. For example, exploring the entrepreneur’s perspective would be a useful way in which to extend Kelly and Hay’s (2003) pioneering study of business angel contracts. Many aspects of the post-investment relationship also require to be examined. Understanding the relational component is one issue. How do the parties cope with adversity? When do business angels find it necessary to assert their rights and how do they do so? (Kelly and Hay, 2003). Quantifying the impact of the value-added contribution of angels on business performance, and the contributions of different types of business angels, is another issue that requires attention. Mason and Harrison’s (2002a) study of investment returns requires corroboration. Meanwhile, adopting new methodological approaches to explore topics that are better understood (e.g., investment decision-making) might provide new insights or challenge existing understanding. Finally, future research needs to have stronger theoretical foundations. Agency theory—the most commonly used theoretical framework - has been shown to have its limitations in a business angel context (Landström, 1992; Kelly and Hay 2003), thus, there is a need for alternative theoretical perspectives.

## NOTES

<sup>1</sup> The term *angel* was coined by Broadway insiders in the early 1900s to describe wealthy theater-goers who made high-risk investments in theatrical productions. Angels invested in these shows primarily for the privilege of rubbing shoulders with the theater personalities that they admired. The term *business angel* was given to those individuals who perform essentially the same function in a business context (Benjamin and Margulis, 2000, p. 5). There is a long tradition of angel investing in businesses (Sohl, 2003). However, this type of business financing has only become significant since the 1950s and 1960s when a lot of the pioneering garage start-ups in Silicon Valley obtained their initial funding from this source.

<sup>2</sup> Anyone who had personally invested in a business start-up which was not their own, excluding stock and mutual funds.

<sup>3</sup> See Hindle and Rushworth (2001) for an international comparison of angel profiles.

<sup>4</sup> Nevertheless, business angels remain largely a phenomenon of Anglo-Saxon countries. For example, in Europe business angels are much more prominent in North West Europe than in Southern Europe. One possible explanation is that in such countries other actors—such as *Impannatori* in the industrial districts of Italy—perform the functions of business angels (Lazzeretti et al., 2004).

<sup>5</sup> This paradox is explained, at least in part, by the contrasts in data availability. There are databases on formal venture capital investments whereas informal venture capital investments go unrecorded. In addition, venture capital fund managers are listed in directories whereas

business angels are invisible. The consequence is that researchers must adopt creative techniques for identifying business angels and getting them to respond to surveys. In practice, much of the research is based—of necessity—on samples of convenience which cannot be tested for representativeness because the population of angels is unknown (Wetzel, 1981). Difficulties in identifying business angels and low response rates because of their desire for privacy results in small sample sizes which restricts the scope for rigorous statistical analysis (Mason and Harrison, 1994, pp. 71–76; Hindle and Rushworth, 2001, pp. 10–11).

<sup>6</sup> Visser and Williams (2001) emphasized that T&T artists are distinguished from “company doctors” who may be called in to turn a business around but do not necessarily invest their own money, and from “corporate raiders” who may, or may not, invest their own money but whose aim is to sell off valuable components of the business as soon as possible.

<sup>7</sup> For a venture capital fund the transactions costs involved in making investments—the time involved in undertaking the evaluation and negotiation of a deal, professional costs and the provision of post-investment support—are both substantial and largely fixed regardless the size of the investment. In “small” investments, these transaction costs represent a significant proportion of the overall investment, making them uneconomic. Business angels are able to make small investments because they do not cost their time in the same way as a venture capital fund managers and their requirement for professional support, for example, from lawyers and accountants, is minimal.

<sup>8</sup> As Gaston (1989b) noted, the financial needs of new and young businesses are not neatly boxed into separate loan and equity categories. Their capital needs frequently shift between these types. Angels make their investments in the form of loans (usually unsecured), loan guarantees, equity and combinations of these types of finance.

<sup>9</sup> Many business angels suffered serious losses in the technology downturn. Those most affected were investors in technology businesses. Many of these businesses failed as a result of market decline or faulty business models. However, business angels also lost out in situations where businesses were able to raise further funding from either their existing venture capital investors or from new investors. In these circumstances, a combination of the inability of angels to provide follow-on funding, the much lower valuation of the subsequent funding compared with the original investment by the angels (“down-rounds”) and their loss of rights as a result of the very onerous terms and conditions under which the venture capital funds invested in down rounds (e.g., liquidation preferences) resulted in a significant dilution in the angel’s investment, often to the extent of rendering it worthless even if the investee company was a going concern. The consequence of this aggressive behavior by venture capital funds has been to create considerable bad feeling between them and the angel community (Mason, 2006).

<sup>10</sup> However, there is a greater chance of a mismatch between the needs of the entrepreneurs and the preferences and value-added skills of potential investors in such regions. Johnstone (2001) noted that in the case of Cape Breton demand for angel finance is concentrated among IT businesses and they want investors to provide marketing and management inputs whereas the investors typically have no knowledge of the sector and so have limited ability to add value.

<sup>11</sup> Benjamin and Margulis (2000, pp. 205–218) provided an example of a due diligence questionnaire.

<sup>12</sup> Riding et al. (1995) quoted one Canadian investor who said that the potential investee business had to pass what was termed “the Toledo test.” That is, if the angel was not willing to spend a weekend in Toledo (a particularly unattractive U.S. city with few diversions) with the principal(s), the investment would not be undertaken. The British equivalent might be “the Luton test” or “the Hull test” (these cities have the dubious privilege of coming out top of the first and second *Crap Town League*: [www.craptowns.com](http://www.craptowns.com)).

<sup>13</sup> Sørheim (2003, p. 357) makes a similar point. "... Experienced business angels in the study emphasize that they are investing in the very early stage in the life cycle of entrepreneurial ventures. Consequently, they must by-and-large depend on the information provided by the entrepreneur or entrepreneurial team, and are therefore very much concerned with [their] perceived trustworthiness.... The investors in this study perceive the creation of some kind of common platform involving shared goals and values as an antecedent for developing trustworthy relationships between entrepreneurs and [themselves]. If this common platform is found to be lacking they reject the opportunity."

<sup>14</sup> In contrast, Sapienza et al. (1996) argued that venture capitalists adopt a "home run" strategy of focusing their attention on likely winners rather than those businesses in their portfolio which are likely to yield little return.

<sup>15</sup> Harrison and Mason (2004) proposed critical incident analysis as an alternative way in which to assess the contribution of investors.

<sup>16</sup> Several of these angel groups have been profiled in the scholarly literature (May and Simmons, 2001; May, 2002; Cerullo and Sommer, 2002; Payne and Mccarty, 2002; May and O'Halloran, 2003).

<sup>17</sup> For example, in Scotland, there are estimated to be, depending on definition, between six and 12 angel groups which invested around £40m in more than 50 companies. The leading syndicates—for example, Archangels and Braveheart—have high visibility, including their own web sites which list their investments, and their investments are reported in the media. Archangels has been operating for about ten years. Its web site lists 20 investments in which they have invested over £30m. In 2002 it invested £1.5m in six new investments and £4.3m in eight follow-on investments. Some of these investments were made as part of syndicated deals involving other angel syndicates and venture capital funds. Braveheart has been operating since 1997. It has 50 members. It has made 22 investments in 17 companies. To put the scale of their investment in some kind of perspective, both Archangels and Braveheart now make more early stage investments in Scotland than any single venture capital fund. Moreover, both syndicates participate in Scottish Enterprise's Co-investment Scheme, underlining their 'institutional' status. Curiously, in England angel syndicates adopt a much lower profile.

## REFERENCES

- ACOST (1990). *The Enterprise Challenge: Overcoming Barriers to Growth in Small Firms*. HMSO, London.
- Acs, Z. J. and Prowse, S. (2001). Angel Investors and the Angel Capital Network. In *Bridging the Entrepreneurial Financing Gap*, J. Whincop (Ed.), pp. 38–53. Ashgate, Aldershot.
- Aernoudt, R. (1999). Business Angels: Should They Fly on Their Own Wings? *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 187–195.
- Atkin, R. and Esiri, M. (1993). *Informal Investment—Investor and Investee Relationships*. Paper to the 16<sup>th</sup> National Small Firms Policy and Research Conference, Nottingham, 17–19 November.
- Benjamin, G. A. and Margulis, J. B. (2000). *Angel Financing: How to Find and Invest in Private Equity*. Wiley, New York.
- Blair, A. (1996). Creating An Informal Investor Syndicate: Personal Experiences Of A Seasoned Informal Investor. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 156–196. Prentice Hall, Hemel Hempstead.

- Blatt, R. and Riding, A. (1996). Where Angels Fear To Tread? Some Lessons from the Canada Opportunities Investment Network. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 75–88. Prentice Hall, Hemel Hempstead.
- Boyns, N., Cox, M., Spires, R. and Hughes, A. (2003). *Research into the Enterprise Investment Scheme and Venture Capital Trusts*. PACEC, Cambridge.
- Brettel, M. (2003). Business Angels In Germany: A Research Note. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 251–268.
- Brush, C. G., Carter, N. M., Greene, P. G., Mart, M. M. and Gatewood, P. (2002). The Role Of Social Capital And Gender In Linking Financial Suppliers And Entrepreneurial Firms: A Framework For Future Research. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 305–323.
- BVCA (2004). *Report on Investment Activity*. British Venture Capital Association, London.
- Bygrave, W. D. and Reynolds, P. D. (2004). *Who Finances Startups In The U.S.? A Comprehensive Survey Of Informal Investors, 1999–2003*. Paper to the Babson-Kaufman Entrepreneurship Research Conference, University of Strathclyde, Glasgow, 3–5 June.
- Bygrave, W. D., Hay, M., Ng, E. and Reynolds, P. (2003). A Study Of Informal Investing In 29 Nations Composing The Global Enterprise Monitor. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 101–116.
- Cerullo, B. and Sommer, B. (2002). Helping Healthcare Entrepreneurs: A Case Study Of Angel Healthcare Investors, LLC. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 325–330.
- Chang, C. K. N., Shipp, S. S. and Wang, A. J. (2002). The Advanced Technology Program: A Public-Private Partnership For Early Stage Technology Development. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 363–370.
- Clarke, J. (1996). Business Introduction Services In The United Kingdom And The Financial Services Act 1986. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 218–226. Prentice Hall, Hemel Hempstead.
- Coveny, P. and Moore, K. (1998). *Business Angels: Securing Start-Up Finance*. Wiley, Chichester.
- Cressy, R. and Olofsson, C. (1997). The Financial Conditions For Swedish SMEs: Survey And Research Agenda. *Small Business Economics* 9: 179–194.
- Duxbury, L., Haines, G. and Riding, A. (1996). A Personality Profile Of Canadian Informal Investors. *Journal of Small Business Management* 24 (2): 44–55.
- Ehrlich, S. B., De Noble, A. F., Moore, T. and Weaver, R. R. (1994). After The Cash Arrives: A Comparative Study Of Venture Capital And Private Investor Involvement In Entrepreneurial Firms. *Journal of Business Venturing* 9: 67–82.
- EVCA (2005). *EVCA Final Performance Figures For 2004*. Press release available at [www.evca.com](http://www.evca.com)
- Farrell, A. E. (1998). *Informal Venture Capital Investment in Atlantic Canada: A Representative View of Angels?* St Mary's University, Halifax, Canada.
- Feeney, L., Johnstone, H. and Riding, A. L. (1998). A Profile Of Informal Investors In Canada: A Comparison Of Canadian And Maritime Investors. Paper to the CCSBE 15<sup>th</sup> annual conference, Halifax.
- Feeney, L., Haines, G. H. and Riding, A. L. (1999). Private Investors' Investment Criteria: Insights From Qualitative Data. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 121–145.
- Freear, J. and Wetzel, W. E. (1990). Who Bankrolls High-Tech Entrepreneurs? *Journal of Business Venturing* 5: 77–89.

- Freear, J., Sohl, J. E. and Wetzel, W. E. Jr. (1992). The Investment Attitudes, Behavior And Characteristics Of High Net Worth Individuals. In *Frontiers of Entrepreneurship Research 1992*, N. C. Churchill, S. Birley, W. D. Bygrave, D. F. Muzyka, C. Wahlbin and W. E. Wetzel, Jr. (Eds), pp. 374–387. Babson College, Babson Park, MA.
- Freear, J., Sohl, J. E. and Wetzel, W. E. Jr. (1994a). Angels And Non-Angels: Are There Differences? *Journal of Business Venturing* 9: 109–123.
- Freear, J., Sohl, J. E. and Wetzel, W. E. Jr. (1994b). The Private Investor Market For Venture Capital. *The Financier* 1 (2): 7–15.
- Freear, J., Sohl, J. E. and Wetzel, W. E. Jr. (1995). Angels: Personal Investors In The Venture Capital Market. *Entrepreneurship and Regional Development* 7: 85–94.
- Fiet, J. O. (1995). Risk Avoidance Strategies In Venture Capital Markets. *Journal of Management Studies* 32: 551–574.
- Fried V. H. and Hisrich R. D. (1994). Toward A Model Of Venture Capital Investment Decision Making. *Financial Management* 23 (3): 28–37.
- Fried, V. H., Bruton, G. D. and Hisrich, R. D. (1998). Strategy And The Board Of Directors In Venture Capital-Backed Firms. *Journal of Business Venturing* 13: 493–503.
- Gaston, R. J. (1989a). The Scale Of Informal Capital Markets. *Small Business Economics* 1: 223–230.
- Gaston, R. J. (1989b). *Finding Venture Capital for Your Firm: A Complete Guide*. John Wiley & Sons, New York.
- Gaston, R. J. (1990). Financing Entrepreneurs: The Anatomy Of A Hidden Market. In *Financing Economic Development*, R. D. Bingham, E. W. Hill and S. B. White (Eds), pp. 266–284. Sage, Newbury Park.
- Haines, G. H., Jr., Madill, J. J. and Riding, A. L. (2003). Informal Investment In Canada: Financing Small Business Growth. *Journal of Small Business and Entrepreneurship* 16 (3/4): 13–40.
- Haar, N. E., Starr, J. and MacMillan, I. C. (1988). Informal Risk Capital Investors: Investment Patterns On The East Coast Of The U.S. *Journal of Business Venturing* 3: 11–29.
- Harrison, R. T. and Mason, C. M. (1992a). International Perspectives On The Supply Of Informal Venture Capital. *Journal of Business Venturing* 7: 459–475.
- Harrison, R. T. and Mason, C. M. (1992b). The Roles Of Investors In Entrepreneurial Companies: A Comparison Of Informal Investors Venture Capitalists. In *Frontiers of Entrepreneurship Research 1992*, N. C. Churchill, S. Birley, W. D. Bygrave, D. F. Muzyka, C. Wahlbin and W. E. Wetzel Jr. (Eds), pp. 388–404. Babson College, Babson Park, MA.
- Harrison, R. T. and Mason, C. M. (1996). Developing The Informal Venture Capital Market: A Review Of The Department Of Trade And Industry's Informal Investment Demonstration Projects. *Regional Studies* 30: 765–771.
- Harrison, R. T. and Mason, C. M. (2000). Venture Capital Market Complementarities: The Links Between Business Angels And Venture Capital Funds In The U.K. *Venture Capital: An International Journal of Entrepreneurial Finance* 2: 223–242.
- Harrison, R. T. and Mason, C. M. (2004). *A Critical Incident Analysis Technique Approach To Entrepreneurial Research: Developing A Methodology To Analyse The Value-Added Contribution Of Informal Investors*. Paper to the Babson-Kauffman Entrepreneurship Research Conference, University of Strathclyde, Glasgow, 3–5 June.
- Harrison, R. T. and Mason, C. M. (2005). *Does Gender Matter? Women Business Angels And The Supply Of Entrepreneurial Finance In The United Kingdom*. Hunter Centre for Entrepreneurship, University of Strathclyde and Centre for Entrepreneurship Research, Edinburgh University Management School, Glasgow and Edinburgh.
- Higashide, H. and Birley, S. (2002). The Consequences Of Conflict Between The Venture Capitalist And The Entrepreneurial Team In The United Kingdom From The Perspective Of The Venture Capitalist. *Journal of Business Venturing* 17: 59–81.

- Hindle, K. and Rushworth, S. (2001). The Demography of Investor Heaven: International Research on the Attitudes, Behaviors and Characteristics of Business Angels. In *Bridging the Entrepreneurial Finance Gap*, M. J. Whincop (Ed.), pp. 10–37. Ashgate, Aldershot.
- Hindle, K. and Wenban, R. (1999). Australia's Informal Venture Capitalists: An Exploratory Profile. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 169–186.
- Hindle, K. and Lee, L. (2002). An Exploratory Investigation Of Informal Capitalists In Singapore. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 169–181.
- HM Treasury (2004a). *Informal Capital Raising and High Net Worth and Sophisticated Investors: A Consultation Document on Proposed Changes to the Financial Promotion Order*. HM Treasury, London.
- HM Treasury (2004b). *Informal Capital Raising and High Net Worth and Sophisticated Investor: Changes to the Financial Promotion Order: Government Response*. HM Treasury, London.
- Industry Canada (2001). *The Winning Formula At Work: Investment Facilitation Techniques*. Industry Canada, Ottawa. Cat No. C2-568/2001E.
- Infometrics Ltd. (2004). *New Zealand's Angel Capital Market: The Supply Side*. Ministry of Economic Development, Wellington.
- Investor Pulse (2003). *U.K. Angel Attitude Survey*. London, C2Ventures Ltd., [www.c2ventures.com](http://www.c2ventures.com).
- Jensen, M. (2002). Angel Investors: Opportunity Among Chaos. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 295–304.
- Jensen, M. C. and Meckling, W. H. (1976). Theory Of The Firm: Managerial Behavior, Agency Costs, And Ownership Structure. *Journal of Financial Economics* 3: 305–360.
- Johnstone, H. (2001). Equity Gaps In Depleted Communities: An Entrepreneurial Response. In *New Technology-Based Firms in the New Millennium*, W. Daring, R. Oakey and S. Kauser (Eds), pp. 84–94. Pergamon, Oxford.
- Kelly, P. and Hay, M. (1996). Serial Investors And Early Stage Finance. *Journal of Entrepreneurial and Small Business Finance* 5: 159–174.
- Kelly, P. and Hay, M. (2000). 'Deal-Makers': Reputation Attracts Quality. *Venture Capital: An International Journal of Entrepreneurial Finance* 2: 183–202.
- Kelly, P. and Hay, M. (2003). Business Angel Contracts: The Influence Of Context. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 287–312.
- Landström, H. (1992). The Relationship Between Private Investors And Small Firms: An Agency Theory Approach, *Entrepreneurship and Regional Development* 4: 199–223.
- Landström, H. (1993). Informal Risk Capital In Sweden And Some International Comparisons, *Journal of Business Venturing* 8: 525–540.
- Landström, H. (1998). Informal Investors As Entrepreneurs. *Technovation* 18: 321–333.
- Lange, J., Leleux, B. and Surlemont, B. (2002). Angel Networks For The 21<sup>st</sup> Century: An Examination Of Practices Of Leading Networks In Europe And The U.S. In *The Growth of Venture Capital: A Cross-Cultural Comparison*, D. Cetindamar (Ed.). Quorum Books, Westport, CT.
- Lazzeretti, L., De Propriis, L. and Storai, D. (2004). *Impannatori* and Business Angels: Two Models of Informal Capital Provision. *International Journal of Urban and Regional Research* 28: 839–854.
- Lengyel, Z. and Gulliford, J. (1997). *The Informal Venture Capital Experience*. Local Investment Networking Company, London.
- Lindsay, N. J. (2004). Do Business Angels Have Entrepreneurial Orientation? *Venture Capital: An International Journal of Entrepreneurial Finance* 6: 197–210.
- Lindström, G. and Olofsson, C. (2001). Early Stage Financing Of NTBFs: An Analysis Of Contributions From Support Actors. *Venture Capital: An International Journal of Entrepreneurial Finance* 3: 151–168.

- Lipper, G. and Sommer, B. (2002). Encouraging Angel Capital: What The US States Are Doing. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 357–361.
- Lumme, A., Mason, C. and Suomi, M. (1998). *Informal Venture Capital: Investors, Investments and Policy Issues in Finland*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Madill, J. J., Haines, G. H. Jr. and Riding, A. L. (2005). The Role Of Angels In Technology SMEs: A Link To Venture Capital. *Venture Capital: An International Journal of Entrepreneurial Finance* 7: 107–129.
- Mason, C. M. (2006). The Informal Venture Capital Market In The United Kingdom: Adding The Time Dimension. In *Research on Entrepreneurship and Management. Volume 5: Venture Capital and Entrepreneurship*, J. Butler (Ed.), pp. 137–171. Information Age Publishing, Greenwich, CT.
- Mason, C. M. and Harrison, R. T. (1993). Strategies For Expanding The Informal Venture Capital Market. *International Small Business Journal* 11 (4): 23–38.
- Mason, C. M. and Harrison, R. T. (1994). The Informal Venture Capital Market In The U.K. In *Financing Small Firms*, A. Hughes and D. J. Storey (Eds), pp. 64–111. Routledge, London.
- Mason, C. M. and Harrison, R. T. (1995). Closing The Regional Equity Gap: The Role Of Informal Venture Capital. *Small Business Economics* 7: 153–172.
- Mason, C. M. and Harrison, R. T. (1996a). Informal investment business introduction services: some operational considerations. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 27–58. Prentice Hall, Hemel Hempstead.
- Mason, C. M. and Harrison, R. T. (1996b). LINC: A Decentralized Approach To The Promotion Of Informal Venture Capital. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 119–141. Prentice Hall, Hemel Hempstead.
- Mason, C. M. and Harrison, R. T. (1996c). Why Business Angels Say No: A Case Study Of Opportunities Rejected By An Informal Investor Syndicate. *International Small Business Journal* 14 (2): 35–51.
- Mason, C. M. and Harrison, R. T. (1996d). Informal Venture Capital: A Study Of The Investment Process And Post-Investment Experience. *Entrepreneurship and Regional Development* 8: 105–126.
- Mason, C. M. and Harrison, R. T. (1997). Business Angel Networks And The Development Of The Informal Venture Capital Market In The U.K.: Is There Still A Role For The Public Sector? *Small Business Economics* 9: 111–123.
- Mason, C. M. and Harrison, R. T. (1999). Public Policy And The Development Of The Informal Venture Capital Market: U.K. Experience And Lessons For Europe. In *Industrial Policy in Europe*, K. Cowling (Ed.), pp. 199–223. Routledge, London.
- Mason, C. M. and Harrison, R. T. (2000a). Informal Venture Capital And The Financing Of Emergent Growth Businesses. In *The Blackwell Handbook of Entrepreneurship*, D. Sexton and H. Landström (Eds), pp. 221–239. Blackwell, Oxford.
- Mason, C. M. and Harrison, R. T. (2000b). The Size Of The Informal Venture Capital Market In The United Kingdom. *Small Business Economics* 15: 137–148.
- Mason, C. M. and Harrison, R. T. (2000c). Influences On The Supply Of Informal Venture Capital In The U.K.: An Exploratory Study Of Investor Attitudes. *International Small Business Journal* 18 (4): 11–28.
- Mason, C. M. and Harrison, R. T. (2001). ‘Investment Readiness’: A Critique Of Government Proposals To Increase The Demand For Venture Capital. *Regional Studies* 34: 663–668.
- Mason, C. M. and Harrison, R. T. (2002a). Barriers To Investment In The Informal Venture Capital Sector. *Entrepreneurship and Regional Development* 14: 271–287.

- Mason, C. M. and Harrison, R. T. (2002b). Is It Worth It? The Rates of Return from Informal Venture Capital Investments. *Journal of Business Venturing* 17: 211–236.
- Mason, C. M. and Harrison, R. T. (2003). ‘Auditioning For Money’: What Do Technology Investors Look For At The Initial Screening Stage? *Journal of Private Equity* 6 (2): 29–42.
- Mason, C. M. and Harrison, R. T. (2004a). Improving Access To Early Stage Venture Capital In Regional Economies: A New Approach To Investment Readiness. *Local Economy* 19: 159–173.
- Mason, C. M. and Harrison, R. T. (2004b). Does Investing In Technology-Based Firms Involve Higher Risk? An Exploratory Study Of The Performance Of Technology And Non-Technology Investments By Business Angels. *Venture Capital: An International Journal of Entrepreneurial Finance* 6: 313–332.
- Mason, C. M. and Harrison, R. T. (2004). New Issues In Venture Capital: An Introduction To The Special Issue. *Venture Capital: An International Journal of Entrepreneurial Finance* 6: 95–103.
- Mason, C. and Rogers, A. (1996). *Understanding the Business Angel’s Investment Decision*. Venture Finance Working Paper No. 14, University of Southampton and University of Ulster, Southampton.
- Mason, C. and Rogers, A. (1997). The Business Angel’s Investment Decision: An Exploratory Analysis. In *Entrepreneurship in the 1990s*, D. Deakins, P. Jennings and C. Mason (Eds), pp. 29–46. Paul Chapman Publishing, London.
- Mason, C. and Stark, M. (2004). What Do Investors Look For In A Business Plan? A Comparison Of The Investment Criteria Of Bankers, Venture Capitalists And Venture Capitalists. *International Small Business Journal* 22: 227–248.
- Mason, C., Cooper, S. and Harrison, R. (2002). Venture Capital And High Technology Clusters: The Case Of Ottawa. In *New Technology-Based Firms in the New Millennium*, Volume II, R. Oakey, W. Daring and S. Kauser (Eds), pp. 261–278. Pergammon, Oxford.
- May, J. (2002). Structured Angel Groups In The U.S.: The Dinner Club Experience. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 337–342.
- May, J. and O’Halloran, E. F. (2003). *Cutting Edge Practices in American Angel Investing*. The Darden School, Batten Institute, University of Virginia, Charlottesville, VA.
- May, J. and Simmons, C. (2001). *Every Business Needs An Angel: Getting The Money You Need To Make Your Business Grow*. Crown Business, New York.
- Murray, G. (1999). Seed Capital Funds And The Effect Of Scale Economies. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 351–384.
- Paul, S., Whittam, G. and Johnston, J. B. (2003). The Operation Of The Informal Venture Capital Market In Scotland. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 313–335.
- Payne, W. H. and Mccarty, M. J. (2002). The Anatomy Of An Angel Investing Network: Tech Coast Angels. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 31–336.
- Pereiro, L. (2002). Tango And Cash: Entrepreneurial Finance And Venture Capital In Argentina. *Venture Capital: An International Journal of Entrepreneurial Finance* 3: 291–308.
- Politis, D. and Landström, H. (2002). Informal Investors As Entrepreneurs—The Development Of An Entrepreneurial Career. *Venture Capital: An International Journal of Entrepreneurial Finance* 4: 77–101.
- Preston, S. L. (2004). *Angel Investment Groups, Networks and Funds: A Guidebook to Developing the Right Angel Organization For Your Community*. Kauffman Foundation, Kansas City.
- Prowse, S. (1998). Angel Investors And The Market For Angel Investments. *Journal of Banking and Finance* 22: 785–792.



- Reitan, B. and Sørheim, R. (2000). The Informal Venture Capital Market In Norway—Investor Characteristics, Behaviors And Investment Preferences. *Venture Capital: An International Journal of Entrepreneurial Finance* 2: 129–141.
- Reynolds, P. D. (1997). Who Starts New Firms? Preliminary Exploration Of Firms-In-Gestation. *Small Business Economics* 9: 449–462.
- Riding, A. L. and Short, D. M. (1989). Some Investor And Entrepreneur Perspectives On The Informal Market For Risk Capital. *Journal of Small Business and Entrepreneurship* 5 (2): 19–30.
- Riding, A. L., Dal Cin, P., Duxbury, L., Haines, G. and Safrata, R. (1993). *Informal Investors in Canada: The Identification of Salient Characteristics*. Carleton University, Ottawa.
- Riding, A. L., Duxbury, L. and Haines, G. Jr. (1995). *Financing enterprise development: decision-making by Canadian angels*. School of Business, Carleton University, Ottawa.
- Sætre, A. S. (2003). Entrepreneurial Perspectives On Informal Venture Capital. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 71–94.
- San José, A., Roure, J. and Aernoudt, R. (2005). Business Angel Academies: Unleashing The Potential For Business Angel Investment. *Venture Capital: An International Journal of Entrepreneurial Finance* 7: 149–165.
- Sapienza, H. J. and Gupta, A. K. (1994). The Impact Of Agency Risks And Task Uncertainty On Venture Capitalist-CEO Interaction. *Academy of Management Journal* 37: 1618–1632.
- Sapienza, H. J., Manigart, S. and Vermeir, W. (1996). Venture Capitalist Governance And Value Added In Four Countries. *Journal of Business Venturing* 11: 439–469.
- Shepherd, D. A. and Zacharakis, A. (1999). Conjoint Analysis: A New Methodological Approach For Researching Venture Capitalists Decision Policies. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 197–217.
- Short, D. M. and Riding, A. L. (1989). Informal Investors In The Ottawa-Carleton Region: Experiences And Expectations. *Entrepreneurship and Regional Development* 1: 99–112.
- Sohl, J. E. (1999). The Early Stage Equity Market In The United States. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 101–120.
- Sohl, J. E. (2003). The Private Equity Market In The U.S.: Lessons From Volatility. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 29–46.
- Sohl, J. E., van Osnabrugge, M. and Robinson, R. J. (2000). *Models Of Angel Investing: Portals To The Early Stage Market*. Paper to the Babson-Kauffman Entrepreneurship Research Conference, Babson College, MA.
- Sørheim, R. (2003). The Pre-Investment Behaviors Of Business Angels: A Social Capital Approach. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 337–364.
- Sørheim, R. (2005). Business angels as facilitators for further finance: an exploratory study. *Journal of Small Business and Enterprise Development* 12: 178–191.
- Sørheim, R. and Landström, H. (2001). Informal Investors—A Categorization With Policy Implications. *Entrepreneurship and Regional Development* 13: 351–370.
- Stedler, H. R. and Peters, H. H. (2003). Business Angels In Germany: An Empirical Study. *Venture Capital: An International Journal of Entrepreneurial Finance* 5: 269–276.
- SQW (2004). *Evaluation of the Investment Readiness Demonstration Projects and Fit4Finance*. SQW Limited, Cambridge.
- Sullivan, M. K. (1994). Altruism And Entrepreneurship. In *Frontiers of Entrepreneurship Research*, W. D. Bygrave, S. Birley, N. C. Churchill, E. Gatewood, F. Hoy, R. H. Keeley and W. E. Wetzel Jr. (Eds), pp. 373–380. Babson College, Babson Park, MA.
- Tashiro, Y. (1999). Business Angels In Japan. *Venture Capital: An International Journal of Entrepreneurial Finance* 1: 259–273.
- Tyebjee, T. and Bruno, A. (1984). A Model Of Venture Capitalist Investment Activity. *Management Science* 30: 1051–1066.

- Tymes, E. R. and Krasner, O. J. (1983). Informal Risk Capital In California. In *Frontiers of Entrepreneurship Research 1983*, J. A. Hornaday, J. A. Timmons and K. H. Vesper (Eds), pp. 347–368. Babson College, Babson Park, MA.
- Van Osnabrugge, M. (2000). A Comparison Of Business Angel And Venture Capitalist Investment Procedures: An Agency Theory-Based Analysis. *Venture Capital: An International Journal of Entrepreneurial Finance* 2: 91–109.
- Visser, R. and Williams, R. (2001). Prospecting For Gold: How Dutch Informal Investors Appraise Small Businesses In Trouble. *Venture Capital: An International Journal of Entrepreneurial Finance* 3: 1–24.
- Wetzel, W. E. (1981). Informal Risk Capital In New England. In *Frontiers of Entrepreneurship Research 1981*, K. H. Vesper (Ed.), pp. 217–245. Babson College, Wellesley, MA.
- Wetzel, W. E. Jr. (1983). Angels And Informal Risk Capital. *Sloan Management Review* 24 (4): 23–34.
- Wetzel, W. E. Jr. (1986). Entrepreneurs, Angels And Economic Renaissance. In *Entrepreneurship, Intrapreneurship and Venture Capital*, R. D. Hisrich (Ed.), pp. 119–139. Lexington Books, Lexington, MA.
- Wetzel, W. E. Jr. (1987). The Informal Risk Capital Market: Aspects Of Scale And Efficiency. *Journal of Business Venturing* 2: 299–313.
- Wetzel, W. E. Jr. and Freear, J. (1996). Promoting Informal Venture Capital In The United States: Reflections On The History Of The Venture Capital Network. In *Informal Venture Capital: Evaluating the Impact of Business Introduction Services*, R. T. Harrison and C. M. Mason (Eds), pp. 61–74. Prentice Hall, Hemel Hempstead.
- Zider, B. (1998). How Venture Capital Works. *Harvard Business Review*, Nov-Dec, 131–139.

JOÃO PEDRO AZEVEDO

*Institute of Applied Economic Research (IPEA), Av. Pres. Antonio Carlos, 51-10º andar,  
Rio de Janeiro-RJ, 20020-010, Brazil*

*E-mail: jazevedo@ipea.gov.br*

## 11. Microfinance and Poor Entrepreneurs

### 1. INTRODUCTION

In the early 1990s, a new financial instrument gained enormous popularity throughout the development community. The promise was to enable the poor to pay their way out of poverty. The flagship was the Grameen Bank, in Bangladesh, which claimed a substantial number of clients, a surprisingly low default rate and almost 20 years of experience; the name: Microcredit.<sup>1</sup> Yielding the highest hopes, the Microcredit Summit of February 1997 pledged, perhaps over-optimistically (Rogaly, 1996; Mosley and Hulme, 1998), to reach, with this instrument alone, 100 million families, that is, one-half of the world's poor before 2005 (Microcredit Summit Report, 1997).<sup>2</sup>

Almost a decade passed and the euphoria of the magic bullet has rightfully diminished (Gulli, 1998). Analysts and practitioners alike are substantially more cautious and aware of both the potential and the limitations of this instrument. As the industry matured, it became apparent that the poor needed more than just credit and that the supply of other financial services, such as savings, deposits and transfers, were also badly needed. The organizations that came to provide these services became known as Microfinance Institutions (MFIs).

The term *microfinance* refers to the provision of financial services for low-income households and micro entrepreneurs (both urban and rural) for productive purposes. These services are delivered through the use of nonstandard methodologies such as character-based lending, group guarantees and short-term repeated loans. The common elements are that the served clients typically lack the characteristics (e.g., titled property as collateral) required by traditional banks or are beyond the reach of the existing financial network, a market

which, to be reached, demands innovative methods and specialized products or institutions. It is worth observing that this definition does not include institutions that solely specialize in consumer credit and salary-backed loans.<sup>3</sup>

The micro entrepreneurs being targeted are small-scale units that produce and distribute goods and services, operating with very little or no capital, that enter the market at their own risk and which, in most cases, originate from strategies for survival (ILO, 1999). This definition includes micro enterprises, in the strict sense of the term, as well as self-employed producers. The initial low level of capital investment, characteristic of these micro entrepreneurs, tends to create a vicious cycle consisting of low labor productivity, low income and meager profits for reinvesting in the enterprise.

In many regions of the world, MFIs are the natural financial link to micro enterprise owners, particularly to those beyond the financial frontier. Although micro entrepreneurs are not necessarily poor,<sup>4</sup> most of the poor tend to be engaged in occupations related to micro enterprises.<sup>5</sup> Thus, even if the credit and other financial products provided by MFIs are not specifically targeted to the poor, their market is necessarily related (directly or indirectly) to the improvement of the productive possibilities of the poor (either as entrepreneurs themselves or as employees of micro enterprises).<sup>6</sup>

In spite of this link with poverty reduction, it is important to note that MFIs cannot solve the problems faced by the poorest of the poor. This group lacks not only the physical capital but, most importantly, the human capital to leverage themselves out of poverty and would therefore need a more comprehensive set of policies to tackle their challenge (Morduch, 1999, p. 1610).

The objective of this chapter is to survey the theoretical and empirical literature on why and how micro entrepreneurs benefit from the provision of microcredit and associated financial services. The next section discusses who the micro entrepreneurs are, in which businesses and industries they operate and how policies targeted to this group can have a pro-poor dimension. Section 3 shows why formal credit markets often fail to serve these entrepreneurs. Section 4 describes the product innovations that characterize the credit operations of microfinance schemes. Section 5 discusses who pays for MFIs, including their funding mechanisms and their financial self-sustainability. Section 6 outlines the main findings from the empirical literature on the impact of MFIs. The final section concludes with considerations and policy implications relating to microfinance as a micro enterprise promotion tool.

## 2. WHO ARE THE POOR ENTREPRENEURS?

This section describes the kinds of businesses and industries in which poor entrepreneurs operate, presents the type of financial products they demand

and explores the relationship between these entrepreneurs and poverty reduction and alleviation policies.

### *2.1. What Businesses and Industries do They Operate?*

The micro enterprises referred to in this study produce both services and manufactured goods consistent with the definition often found in the literature (ILO, 1999; Paetkau, 1999; Schreiner, 2004b). Most of these are small firms or subsistence farms. These micro enterprises are usually owned by an individual, perhaps supported by one or two family members, engaged in an income-generating activity. It is likely to be a trading activity, selling food or clothing in street stalls or making and selling handicrafts up-country. More often than not, there is very little investment in fixed assets, inventory investment is usually minimal, and the business frequently operates on a cash-only basis as there may be no trade credit extended by suppliers. Enterprise sales may not be distinguished from any other kind of income the individual or household earn. These “enterprises” may be quite unstable and footloose. If better opportunities arise or if costs cannot be covered, the individuals involved might probably move on. They often endure a “hand-to-mouth” existence.

These micro entrepreneurs evidently differ from their counterparts in developed economies which usually offer nontradable services such as child care, haircuts, retail sales, transportation, or home, car or office maintenance. Another important difference is that in the developed economies the average person does not spend a large share of his or her disposable income on purchases from micro enterprises.

### *2.2. What Financial Products do They Demand?*

Micro enterprises have diverse demands for credit because they have diverse cash flows and diverse investment goals. For example, not all demands are satisfied by high-volume short-term working capital through joint-liability loans. The demand for microfinance by these micro enterprises can be classified along a continuum per enterprise size. Income-generating activities (IGAs) appear at the smaller end, micro enterprises in the middle and small enterprises at the larger end (Waterfield, 1994).

IGAs are part-time, often seasonal economic activities undertaken by semi-subsistence households. IGAs generally employ only the entrepreneur and have less than \$500 in assets. With many diversified sources of income, the entrepreneur consumes the profits of the IGA. IGAs demand small short-term working-capital loans and, therefore, they may be able and willing to pay high interest rates.

*Micro enterprises* are larger than IGAs, having less than 10 employees and less than \$10,000 in assets. The micro enterprise is the family's chief source of income and profits are partly reinvested and partly consumed. Micro enterprises demand more flexible financial products than IGAs because micro enterprises seek to finance both working capital and fixed assets. Investments in fixed assets mean larger loans with longer terms and, as a consequence, borrowers seek interest rates lower than those levied on IGAs.

*Small enterprises* are larger than micro enterprises, with less than 50 employees and less than \$100,000 in assets. A small enterprise is the owner's primary source of income and most profits are reinvested for further growth. Small enterprises demand larger, longer loans and are best served by banks. Like other bank borrowers, small enterprises can offer physical collateral and can afford commercial interest rates.

Waterfield's continuum by size is not the only useful typology of demand for microfinance by micro enterprises. For example, subsector analysis of demand is appropriate because different types of micro enterprises operate in different economic markets. Product design may also vary by sector (commerce, manufacturing or service), location (rural or urban), purpose (fixed assets, working capital, consumption or some combination), production type (continuous assembly or made-to-order), gender (man or woman) and borrower experience (new or repeat).

### 2.3. *What are the Links with Poverty Reduction?*

The only permanent way to escape poverty and steadily improve long-term well being is to build assets, be they physical (land or homes), financial (bank accounts), social (networks) or human (education and experience). For many of the world's poorest, the micro enterprise is a way to build assets (Schreiner, 2004a).

Aid for asset-building through micro enterprise has focused on loans (Daley-Harris, 2003) but has also included savings services (Hickson, 2001), grants (Pretes, 2002) and training (Edgcomb, 2002). Although training and grants are not usually delivered by MFIs (with the exception of those institutions with a more integrated approach), they often deliver loans, savings and other financial services.

What keeps the poor poor is their meager human and complementary capital. With income barely above subsistence, they can save little and thus build few assets. With few assets, they lack tools or land that would improve the productivity of their human capital. With low productivity, they must start to work young and are unable to invest time in education to increase human capital for themselves and their children. This can create intergenerational poverty persistence. The need to work long and hard also takes a toll on their health,

further decreasing productivity. With few assets to share, the poor have poor friends, so their social networks are less productive. All of these factors keep their incomes low.

Hence, micro enterprise programs in less-developed and developing countries, such as the promotion of MFIs, can help expand the productive capabilities of the working poor and can be part of a broader poverty alleviation and reduction strategy.

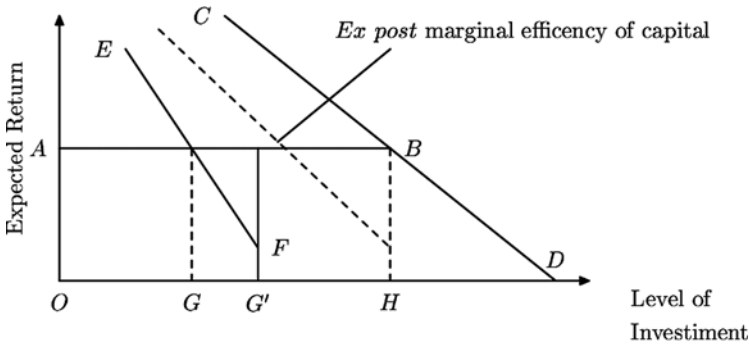
### 3. WHY DO FORMAL CREDIT MARKETS FAIL POOR ENTREPRENEURS?

Credit markets are considered to *fail* if poor entrepreneurs are unable to borrow for socially beneficial projects, that is, projects with an excess of social benefits over social costs. According to Hulme and Mosley (1996), this can happen for any one of four reasons:

1. No lender is willing (or legally permitted) to pass on the extra costs associated with lending in small quantities to unknown customers.
2. No insurer is willing to compensate for borrowers' and lenders' risk-aversion (and for the presumed absence of collateral) by offering insurance against non-repayment due to natural hazards.
3. Even if 1 and 2 do not hold, potential borrowers are unwilling to borrow because of risk aversion, although the expected value of their profits outweighs the expected cost of their investment, including interest and insurance payments.
4. The social and private values of cost and benefit diverge because of externalities so that some projects which are socially profitable do not generate private returns that exceed their costs.

Figure 11-1 illustrates this. CD and EF represent the expected profitability (marginal efficiency) of capital for groups of micro entrepreneurs who are, respectively, risk neutral and risk averse. AB is the cost curve associated with lending to small-scale entrepreneurs, taking into account the additional cost of administration and insurance when dealing with the latter groups. All potential borrowers with projects in regions OG will invest, all borrowers with projects in region DH will not and those in region GH will either invest or not according to risk aversion and the size of business. In Figure 11-1, all those in region GG' fail to invest even though loan and insurance markets (by hypothesis) exist and their investment, if carried out, would have been profitable. This is case 3 above.

It is now possible to consider the role of collateral and insurance. A lender who raises the value of the collateral requirement from zero to some



Source: Hulme and Mosley (1996)

FIGURE 11-1 Risk, cost of capital and market failure.

positive amount raises the expected return  $E$  to him or herself:

$$E(X) = rX(1 - \pi) + (C - X - rX)\pi, \tag{1}$$

where  $\pi$  is the probability of failure of the project and  $C$  the value of the collateral to the lender,  $r$  is the rate of interest paid by the borrower and  $X$ , the amount of the loan. As collateral is increased from zero, the second term (the lender’s pay-off if the project fails) increases progressively since only the difference between the collateral and the loan amount plus interest is lost. If (as is true with most poor entrepreneurs) the borrower is too poor to offer collateral, the alternative for the lender is to add an insurance premium to the interest rate  $r$ , either in the form of a compulsory savings requirement or by requiring the borrower to subscribe to a private insurance scheme; in this event, the insurance scheme pays out an indemnity (in place of collateral  $C$ ) if the project fails. A risk-averse lender, by raising the collateral (or insurance) requirement to levels above the loan size, can cause the expected return to exceed the interest rate, a technique that can be used to circumvent the effect of interest rate ceilings which still exist in several developing countries.

Hulme and Mosley (1996) show that insurance markets may fail to exist for the riskiest (and generally poorest) borrowers. In only one of the 13 microfinance schemes analyzed by the authors (namely., Indian regional rural banks) did borrowers have automatic access to an insurance scheme external to the lender and only four others (namely., Indonesia’s KUPEDDES and BKK, Bangladesh’s Grameen Bank and Kenya’s KREP) had a scheme of insurance for project failure out of savings compulsorily deposited by borrowers. At the high-risk end of the markets it is common for insurance to fail to exist because of the potential insurer’s inability to gauge the size of the risk they are subject to, compounded by “adverse selection,” that is, the tendency of bad risks to



demand insurance and of good risks not to demand it. A vicious circle is easily developed: the insurer's miscalculation of the true risk first time round leads it not to raise the insurance premium to a "realistic" level, but simply to withdraw from any attempt to insure what it perceives as the higher-risk individuals. Then for some would-be insurance purchasers (poor borrowers in Figure 11-1) the cost of borrowing-plus-insurance does not just increase above AB, but effectively becomes infinite: a service which, irrespective of price, they cannot buy because it is not offered. Under these adverse circumstances, they will drop out of the market for loanable funds and not invest. This is case 2 of market failure above.

Another problem is that lenders usually have no reference as to borrowers' intentions about paying back loans. A utility-maximizing borrower, using the notation of (1), will default if the utility of his or her wealth,  $W$ , less loss of future earnings from default ( $D$ ) and of collateral, if any is taken ( $C$ ), exceeds the utility of that wealth when the loan amount plus interest is repaid, that is:

$$U(W - D - C) < U(W - X - rX). \quad (2)$$

Hence, each lender must make a personal judgment about whether (2) is positive or negative for each potential borrower and, if negative, refuse to lend, irrespective of price. Inability to take collateral ( $C = 0$ ) and legal restrictions on the interest rate lenders are allowed to charge ( $r < \bar{r}$ ) clearly increase the likelihood that a lender's judgment of a borrower's creditworthiness will be adverse. Even if a collateral-poor borrower is willing to pay a higher interest rate to cover the higher risk involved, a lender may still refuse to lend. The utility gain for the borrower from default increases with the interest rate charged, whereas the expected returns to the lender can decrease as the interest rate increases (Stiglitz and Weiss, 1981). This is case 1 of market failure above.

Finally, lenders may fail to lend to some projects because they are socially beneficial but not privately profitable. Two important reasons why this may occur are: first, a discrepancy between market prices and "shadow prices" and, second, the inability of lenders to capture the final value of benefits which they confer by means of, for example, training, research or the provision of information about borrowers to other lenders. This is case 4 of market failure.

For these reasons—externalities, risk aversion among borrowers, inability of insurers to combat adverse selection and inability of lenders to combat moral hazards—"good" but poor entrepreneurs in developing countries may be unable to borrow.

#### 4. CAN MICROFINANCE INSTITUTIONS HELP?

MFIs gained most of their international and domestic donor support based on the claim that they could help to expand the financial frontier by providing credit and other financial services to individuals traditionally excluded from these markets in a sustainable manner. This section presents the main product innovations proposed by these institutions. In addition it describes these institutions and shows how relevant they are in financial markets in different regions of the world.

##### *4.1. How can They Help?*

*4.1.1. Financial Products* According to Hoff and Stiglitz (1990), the mechanisms used by MFIs to protect themselves against bad loans can be classified as direct or indirect methods. Direct methods are those in which the lender expands resources directly in the form of administrative expenses on screening, enforcement or insurance and hopes to gain a more than proportionate reward in the form of a lower default rate. Indirect methods are those in which the lender provides borrowers with an incentive to take actions that will return information on the risk of lending to them and/or that reduce the risk of default.

##### *4.1.1.1. Direct Methods: Intensive Loan Collection and “Incentives to Repay.”*

An important mechanism for securing high repayment rates with high monitoring costs involves exploiting dynamic incentives (Besley and Coate, 1995, p. 2187). Programs typically begin by lending just small amounts and then increasing loan size upon satisfactory repayment. The repeated nature of the interactions—and the credible threat to cut off any future lending when loans are not repaid—can be exploited to overcome information problems and improve efficiency. However, one should bear in mind that these dynamic incentives also run into the problems common to all finite repeated games. If the lending relationship has a clear end, the borrowers have incentives to default in the final period (Morduch, 1999, p. 1583). Anticipating that, the lender will not lend in the final period, giving borrowers incentives to default in the penultimate period—and the game goes on until the entire mechanism is unraveled. Unless there is substantial uncertainty about the end date—or if “graduation” from one program to the next is well established, dynamic incentives have limited scope on their own.

An additional advantage of progressive lending is the ability to test borrowers with small loans at the start (Morduch, 1999, p. 1583). This feature allows lenders to develop relationships with clients over time and to screen out the worst prospects before expanding their loan scale.

Another interesting feature of microfinance credit contracts is that repayments must start nearly immediately after disbursement (Morduch, 1999, p. 1584). The advantages of such schemes are various. Regular repayment schedules screen out undisciplined borrowers and allow the bank to get hold of cash flows before they are consumed or otherwise diverted (Rutherford, 2001). In addition, because the repayment process begins before the investment bears fruit, this means that the household must have an additional income source on which to rely. To insist on weekly repayments, therefore, means that the bank is actually lending partly against the household's steady, diversified income stream, not just on the risky project.

In either instance, the bank's costs are increased by this kind of arrangement and the key operational question for the bank is whether the additional costs—be they the supervisory labor required to monitor such a scheme or the discounts for prompt repayment—will bring about a more than proportionate reduction in default rates and thereby improve its cash flow (Hulme and Mosley, 1996).

*4.1.1.2. Direct Methods: Savings Schemes and Loan Insurance.* While only a handful of programs require collateral, many have substitutes. For example, programs following the Grameen model require that borrowers contribute of 0.5 percent of every unit borrowed to an "emergency fund." The proceeds of such compulsory saving schemes can be used to insure against certain named events which may cause the borrower's project to fail and the borrower to default. Some authors argue that the existence of a savings scheme lowers the probability of default because borrowers feel more obligated to repay loans when they perceive the borrowed money as belonging to their neighbors and colleagues rather than to an impersonal institution (Von Pischke, 1991, p. 311).

*4.1.1.3. Indirect Methods: Group Lending and Other Peer-Monitoring Devices.* An alternative set of strategies which a lender may use is to set up a structure through which the job of monitoring and encouraging prompt loan repayment is partly delegated to borrowers rather than being carried out by the lender alone, that is, indirect reduction of risk (Hulme and Mosley, 1996).

The involvement of members of the local administration in the screening process and the public exposure of defaulters come into this category, but by far the most important indirect lender strategy is to lend to groups of borrowers rather than to individuals. The logic of doing so is twofold. First, administrative costs are reduced for the borrowers; second, the probability of default is reduced as a consequence of peer pressure to repay loans exerted by group members. This peer pressure arises because no member of the group can receive credit as long as any member is in default on a loan. Ghatak (1999) shows how this sorting process can be instrumental in improving repayment

rates resulting in lower interest rates and greater social welfare. He argues that a group lending contract provides a way to price discrimination that is impossible with an individual-lending contract. Ghatak (2000) and Lensink, Gangopadhyay and Ghatak (2005) also show that, by exploiting local information, joint liability lending can improve efficiency compared to standard debt contracts in the presence of asymmetric information about borrower types. Hence, when other potential screening instruments, such as collateral, are not available, joint liability lending becomes a particularly attractive method of lending.

However, it is important to note that Lensink, Hermes and Mehrteab (2005) found that the efficiency of group lending appears to be highly correlated with the characteristics of the group leader. In addition, Lensink, Hermes and Mehrteab also show that peer monitoring by and the social ties with the group leader may help to reduce morally hazardous behavior by group members. In particular, their results indicate that regular contact and short distances between the group leader and group members helps to reduce misuse of loans by individual group members. Moreover, if the group leader knows the other group members before forming the group and if he/she did not change groups in the past, this reduces the probability of moral hazard from occurring in the group. It is important to note that the authors find no link between peer monitoring and social ties on the one hand and the occurrence of moral hazard on the other among the other group members.

In addition, group lending can also play an important role in relation to status verification and repayment enforcement, since group members are often on a much better position than banks to learn about each other's ventures outcomes and behavior. As a consequence, joint liability contracts can encourage group members to prevent partners from defaulting opportunistically and have a positive impact on the mitigation of adverse selection (Parker, 2004, pp. 167–169). By favorably altering the pool of borrowers, group lending contracts can create an incentive for borrowers to sort into groups in which only individuals with the highest chances of success are allowed as members.

Two problems with the group concept are immediately apparent. The first is that considerations of minimizing administrative costs favor large groups, while considerations of minimizing default favor small groups which thereby have a better chance of effectively monitoring one another's behavior. This flags the second problem. While groups do generate the positive effect of peer monitoring, they can also generate the perverse effect that group members who are tempted to default may ask other group members to pay on their behalf, in addition to their own share, an option not available in individual lending schemes.

Another important issue related to group lending is whether, in terms of individual risk profiles, groups will be homogeneous or heterogeneous. Since safe types are always preferred as partners (because their probability of failure

is lower), the question becomes: will the riskier types be willing to make a large enough transfer to the safer types so that both risky and safe do better together? Ghatak and Guinnane (1999) show that there is no mutually beneficial way for risky and safe types to group together. Group lending thus leads to assortative matching (Becker, 1991). In one of the few empirical papers on this subject, Sadoulet and Carpenter (2001) tested whether, as is widely assumed in the theoretical literature, credit groups form homogeneously. Using survey data from Guatemala, the authors consistently find evidence that, even accounting for matching frictions, some borrowers choose to form heterogeneous groups, even when project returns appear independent. This result suggests that borrowers are able to use joint liability to create credible insurance arrangements among group members in missing insurance market environments. Nonetheless, the authors also found that beyond a certain risk level, borrowers are unable to remunerate safer partners for the asymmetric insurance, being thus left to match homogeneously. Lensink and Mehrteab (2003) in a study conducted among MFI group members in Eritrea, also found that microcredit groups are formed heterogeneously.

*4.1.1.4. Other Financial Services: Voluntary Savings.* An additional means of promoting household welfare is the development of mechanisms for safe but liquid savings deposits. Early microfinance organizations, with the exception of the Unit Desa program in Indonesia, were not effective in mobilizing savings and showed little interest in doing so. More recently some microfinance institutions have shown that if appealing interest rates, a conveniently located facility and flexible accounts are offered, even poor households are willing to save. A study of the expansion of rural banking in Mexico clearly shows this possibility. Aportela (2000) measures the impact on savings rates of the expansion of Pahnal, a Mexican savings institute targeted at low-income clients. Aportela uses a difference-in-difference methodology to estimate impacts, finding that an expansion of program availability pushed up savings rates by almost five percentage points—and by almost seven percentage points from some of the poorest households. Moreover, Aportela finds little evidence of the crowding out of other savings instruments, suggesting that much of the savings is due to new savings. Savings mobilizations can be justified for a number of reasons: (i) provision of relatively inexpensive source of capital for re-lending; (ii) creation of a natural pool of clients; and (iii) building up savings enables households to reduce their consumption volatility over time. On the other hand, it may become prohibitively expensive to handle a lot of small deposits accounts.

*4.1.1.5. Other Financial Services: Transfers.* In 2002, remittances to developing countries were estimated at \$80 billion, which is twice the amount of

aid provided by rich countries (World Bank, 2003, p. 198).<sup>7</sup> This money is the fruit of the work of immigrants and plays an increasing role in the financing of development in the workers' home countries. Although not undisputed (Chami, Fullenkamp and Jahjah, 2005), many scholars and practitioners argue that remittances can therefore play a key role in private-sector development efforts, enabling families to receive needed capital for, for example, education, housing and small business start-ups and expansion (Kapur, 2004). Nevertheless, transaction costs can be as much as 10 to 15 percent high even for flows to large, urban markets.

In certain countries, such as Jamaica, there is evidence that low-income households are more likely to receive remittance transfers than high-income ones (CGCED, 2002, p. 75). At least potentially, remittances might act as a self-insurance mechanism since they are uncorrelated with income from other sources, so they can play a significant role in the consumption smoothing of households. Other studies suggest that remittances are responsible for almost 20% of the capital invested in micro enterprises throughout urban Mexico, an additional cumulative investment capital among the studied enterprises of \$1.85 billion (Woodruff and Zenteno, 2001). More recently, several microfinance institutions, in particular credit unions, have been pushing regulatory reforms in the Latin America and the Caribbean region in order to be allowed to serve as intermediaries in the transfer of these resources. This would have a direct impact on their clients and their entry into this market would foster the necessary competition to lower the transfer fees, which in some countries add up to almost 20% of the total amount of remittances (MIF, 2002).

*4.1.2. Non-Financial Products* In order to achieve their poverty reduction mission, some microfinance institutions have chosen a more integrated approach toward their clients (Ledgerwood, 1999, Chapter 3) which include several nonfinancial components such as education, business development services and health, as opposed to institutions that have chosen a minimalist approach, focusing solely on the provision of financial services.

Smith (2002) uses some quasi-experimental data from urban Honduras and rural Ecuador to compare the impact of tying-in micro enterprise credit with other services, notably health and education, in village bank programs. The paper presents evidence that in Honduras "health" bank participation is robustly associated with the subsequent reduction of the probability of child diarrhea, but in no specification does credit-only bank participation have this effect. In Ecuador, results suggest a larger effect of credit-only banks. In both countries, health bank participation significantly raises subsequent healthcare over credit-only participation. The authors finally conclude that although further work is needed, it is clear that credit tie-in programs cannot be summarily dismissed as an unproductive interference on the natural comparative advantage

of institutions designed to provide credit to the poor. These results corroborate McKernan's (2002) findings that show non-credit aspects to be an important determinant of the success of micro-credit programs, having large positive effects on self-employment profits.

In his study, Mosley (2001) concludes that there is little difference in poverty impact and in recovery rates between the "minimalist" approaches of BancoSol and PRODEM and the "integrated" principle of ProMujer and SARTAWI, in which credit is supplied in conjunction with additional inputs such as training, health and agricultural services. That does not mean these programs do not help these entrepreneurs, especially because each set of institutions has very different target groups, the latter being poorer than the former.

#### *4.2. Who are They?*

In terms of institutional design, the MFIs form anything but a homogeneous group. Most of this diversity sprang from the grass roots characteristic of these institutions, ranging from informal Rotating Savings and Credit Associations (RoSCAs) and money lenders to formal commercial banks, credit unions and Non-Governmental Organizations (NGOs). Central to the debate is the fact that different institutional forms have different roles in the financial system and furnish different lessons concerning institutional sustainability.

Moneylenders and RoSCAs are informal financial institutions whose use by poor entrepreneurs offers insights into the design of formal institutions (Schreiner, 2001). Moneylenders charge high interest rates but disburse funds quickly and without burdening the borrower with transaction costs (Waterfield, 1994). The use of moneylenders by micro enterprises points to the feasibility of informal guarantees and shows that the poor can pay high interest rates. Moneylenders also indicate the potential for self-sustainability in microfinance. In addition, some authors have claimed that moneylenders can provide a screening function for the formal financial sector, since it often has better information about borrowers (Jain, 1999).

RoSCAs are groups where individuals agree to contribute at regular intervals to a pot that is distributed after each collection to a single group member according to a rule of rotation (Ardener, 1964). RoSCAs enable savings accumulation, transformation of size, limited financial intermediation and insurance (Krahn and Schmidt, 1994).

Commercial banks are slowly tapping into microfinance. Traditional banks have comparative advantages in issuing large loans, contracts supported by the legal system, and transaction accounts (Waterfield, 1994). Some of the important lessons provided by banks are: the importance of economies of scale; a focus on cost-recovery; strong information-management systems;

the development of key ratios for financial analysis; an example of how and when to use the legal system; the importance of external regulations to protect depositors; and the value of different products to meet different demands. The chief comparative advantage of banks—lending based on individual analysis of the applicant’s character, project and past relationship with the bank—is exactly the challenge facing microfinance initiatives. Applying for a bank loan entails high transactions costs for the self-employed poor because they cannot cheaply assemble the loan guarantees, such as physical collateral or enterprise accounting records, that make individual analysis practical (Branch, 1994). High transaction costs, coupled with small loan sizes, ensure that micro enterprises and commercial banks rarely intersect.

Because moneylenders, RoSCAs and commercial banks tend to be poorly suited for micro enterprises, policy makers and practitioners choose NGOs or credit unions to promote and support them. NGOs are private institutions that provide public goods not supplied by governments. In some respects, NGOs are ideal channels to improve financial services for micro enterprise (IPC, 1994), especially in the aspect that they neither are lethargic governments, nor are they profit-seeking banks.

Most NGOs, however, have various severe weaknesses as potentially self-sustainable partners. NGOs are unregulated, often lack financial expertise and usually have weak governance structures. Fed from birth by a steady stream of donor funds, NGOs may be reluctant and ill-equipped to be weaned to self-sustainability. Finally, the charitable origins of many NGOs may make it difficult for them to unlearn paternalistic, nonprofit attitudes.

Credit unions have several comparative advantages as providers of microfinance for micro enterprises (Branch, 1994). They may not require formal collateral beyond a member’s share capital and the share capital of cosigners. This means low transaction costs and allows for small, short-term loans. The credit union’s local membership and its face-to-face relationships allow for flexible loan contracts and the use of local information and peer pressure in screening and enforcement. Best of all, credit unions, when left alone, will fund loans with member deposits and share capital.

#### *4.3. How Relevant are They?*

The latest big picture of the microfinance field was published in 2004 (CGAP, 2004). This study surveyed the global outreach of over 3000 financial institutions all of which focused to some degree on extending financial services downward from the economic level of the traditional clients of commercial banks. These institutions therefore have a “double bottom line”: in addition to a financial objective, they also have a developmental or social objective.



TABLE 11-1 *Active loans<sup>a</sup> in alternative financial institutions (in thousands)*

| Region               | MFIs <sup>b</sup> | Co-ops<br>and<br>credit<br>unions | Rural<br>banks | State/<br>agricultural/<br>development<br>banks | Postal<br>banks | Total         | Percent<br>of total |
|----------------------|-------------------|-----------------------------------|----------------|---|-----------------|---------------|---------------------|
| AFR                  | 3956              | 857                               | 33             | 348   | —               | 5193          | 3%                  |
| EAP<br>(incl. China) | 18,292            | 1069                              | 3147           | 65,624  | —               | 88,133        | 58%                 |
| <i>China only</i>    | <i>153</i>        | <i>18</i>                         | <i>0</i>       | <i>46,570</i>                                   | —               | <i>46,741</i> | <i>31%</i>          |
| ECA                  | 430               | 90                                | —              | 28  | —               | 548           | 0%                  |
| LAC                  | 4464              | 655                               | 162            | 51  | —               | 5332          | 4%                  |
| MENA                 | 909               | 11                                | —              | 5912  | —               | 6832          | 4%                  |
| SA<br>(incl. India)  | 22,366            | 355                               | 1467           | 22,030  | —               | 46,217        | 30%                 |
| <i>India only</i>    | <i>3961</i>       | <i>51</i>                         | —              | <i>19,748</i>                                   | —               | <i>23,760</i> | <i>16%</i>          |
| TOTAL                | 50,415            | 3037                              | 4809           | 93,994  | —               | 152,255       | 100%                |
| %                    | 33%               | 2%                                | 3%             | 62%   | —               | 100%          |                     |

<sup>a</sup>Typically includes loan amounts that have been disbursed, but not repaid or written off.

<sup>b</sup>Includes NGOs, banks, and non-bank financial institutions that specialize in microfinance, as well as microfinance programs in full-service commercial banks.

Key: AFR Africa (sub-Saharan); EAP East Asia and the Pacific; ECA Europe and Central Asia; LAC Latin America and the Caribbean; MENA Middle East and North Africa; SA South Asia.

Source: CGAP (2004).

TABLE 11-2 *Savings accounts in alternative financial institutions (in thousands)*

| Region               | MFIs        | Co-ops<br>and<br>credit<br>unions | Rural<br>banks | State/<br>agricultural/<br>development<br>banks | Postal<br>banks | Total          | Percent<br>of total |
|----------------------|-------------|-----------------------------------|----------------|---|-----------------|----------------|---------------------|
| AFR                  | 3958        | 5648                              | 1113           | 343   | 12,854          | 23,915         | 4%                  |
| EAP<br>(incl. China) | 78,708      | 12,130                            | 6,019          | 15,772  | 141,005         | 253,634        | 44%                 |
| <i>China only</i>    | <i>15</i>   | <i>200</i>                        | —              | —   | <i>110,000</i>  | <i>110,215</i> | <i>19%</i>          |
| ECA                  | 163         | 5691                              | —              | —   | 11,503          | 17,357         | 3%                  |
| LAC                  | 1298        | 8466                              | 48             | 50  | 179             | 10,041         | 2%                  |
| MENA                 | 713         | —                                 | —              | 29  | 16,525          | 46,230         | 8%                  |
| SA<br>(incl. India)  | 18,728      | 1620                              | 11,495         | 53,773  | 136,383         | 221,999        | 9%                  |
| <i>India only</i>    | <i>3927</i> | <i>389</i>                        | —              | <i>50,021</i>                                   | <i>124,010</i>  | <i>178,347</i> | <i>31%</i>          |
| TOTAL                | 103,568     | 33,553                            | 18,675         | 98,930  | 318,450         | 573,176        | 100%                |
| %                    | 18%         | 6%                                | 3%             | 17%   | 56%             | 100%           |                     |

For constituents of MFIs and a key to abbreviations, see previous table.

Tables 11-1 and 11-2 summarize numbers of loans and savings accounts, respectively, by type of institution and geographical region.

The worldwide numbers are dominated by Asia, which has over four-fifths of all MFI accounts, both savings and loans. The Asian numbers, in turn, are dominated by China and India, where there has been a heavy, although not always efficient, government commitment to extend financial services. In terms of Microfinance, the numbers are dominated by Bangladesh and Indonesia, in particular due to the activity of the Grameen Bank and the Unit Desa, respectively.

It may not be surprising that MFIs constitute only a small portion of savings accounts since many MFIs are credit-only institutions. When one looks only at credit services, *MFIs account for about 33% of the loans* (25% of the loans are from NGOs). The MFI share rises to 57% if China and India are excluded. This is an impressive accomplishment given that MFIs as a group are younger and smaller than the other financial institutions covered by the study. In regional terms, Latin America, Eastern Europe and Central Asia, as well as Africa are the three regions with the highest share of MFIs in the active total loan portfolio with 84%, 78% and 76%, respectively.

Although one could not claim these MFIs only finance micro entrepreneurs, it is reasonable to assume that most of these institutions provide productive credit for income generating activities and micro enterprises in developing countries.

## 5. WHO PAYS FOR MICROFINANCE INSTITUTIONS?

International and domestic donors have played a crucial role in the creation and development of MFIs. However, as a few institutions become self-sustainable (Stephens, 2005), their funding possibilities are starting to broaden as awareness of this market by private investors and other commercial financial institutions increases.

For institutions in general, *self-sustainability* means the ability to meet long-term goals. For financial institutions, *self-sustainability* means acting in a way that ensures the return on equity, net of any subsidy, exceeds the opportunity cost of funds. This does not imply pure profit maximization because donors insist not only on self-sustainability but also on the depth of outreach.

This section describes the types of institutions that traditionally fund MFIs around the world; discusses to which extension these institutions are sustainable; and explores the main hurdles for these institutions to reach sustainability.

### 5.1. Who Funds Them?

Donors have played an important role in the development of MFIs. In virtually all cases, the support of these institutions is part of a broader

intervention strategy designed to reduce poverty, generate income, promote employment and/or support entrepreneurship.

Most donors do not have a focal point or a central unit for microfinance (CGAP, 1997). Among the bilateral agencies, the United States Agency for International Development (USAID) and the Department for International Development (DFID) of the U.K. are the only ones with central offices for microfinance, although several are creating similar focal units. Among the regional and multilateral institutions, the Inter-American Development Bank (IDB), the UN Capital Development Fund (UNCDF) and the UN Development Program (UNDP) by way of the Special Unit for Microfinance (SUM) office, and the African Development Bank's AMINA Program have central units dedicated to microfinance. The International Fund for Agriculture Development (IFAD), the UN Conference on Trade and Development (UNCTAD) and the International Labor Office (ILO) reported focal points within each regional office at headquarters as well as in the technical divisions.

While almost all bilateral donors work globally, several donors have an emphasis on Africa, such as the Scandinavian donors (Denmark, Finland, Sweden), France (primarily West Africa) and the U.K. (DFID). Those with a focus on Asia include AusAID (Australia) and DFID. Aside from the Asian, African and Inter-American Regional Development Banks, which have clear geographic priorities, all the multilateral development institutions work globally. Of these, UNCDF/UNDP gives high priority to Africa with 75% of resources targeted for the region.

The majority of bilateral agencies provide grants and/or soft loans and, in limited cases, equity to MFIs. Regional and multilateral development banks typically provide loans to governments, although several have opened microfinance facilities for direct financing of MFIs, such as UNDP's Micro-Start Program, and the CGAP Secretariat at the World Bank.

As MFIs mature and start to become sustainable, a new type of funding possibility emerges, namely, private or institutional funds. Today, 60 listed funds invest in microfinance. These institutions provide grants, equity investments, guarantees and technical assistance to MFIs. Some of these institutions have a particular region of interest or a theme, while others have a broader scope. Among these funds, two deserve special attention.

The first one is the ProCredit Holding (PCH) comprising a specialized target—group-oriented financial institutions that focus on providing business loans to micro-, small- and medium-sized enterprises (MSMEs or SMEs) in developing countries and transition economies. The PCH network has 19 members worldwide (as of 2004-09-30), most of them located in Africa, Eastern Europe and Central Asia, and Latin America and the Caribbean. The institution portfolio invested in MFI comprised of \$89,181,767 (€71,820,000) as of September 2004.

PCH is a development oriented private investment company based in Frankfurt, Germany. It acquires equity stakes in partner commercial financial institutions that are target-group oriented. Its aim is to strengthen them financially and support them in their development orientation by establishing and running target group oriented institutions which “extend the frontier of finance,” that is, provide access to commercial banking services to segments of the population previously neglected by the formal financial sector. In particular the institutions in the PCH group focus on providing credit to micro-, small- and medium-sized enterprises because they believe these enterprises are able to create the largest numbers of jobs and make a vital contribution to the economies in which they operate.

The second important player in this field is the International Financial Corporation (IFC), a multilateral organization that is part of the World Bank Group. One of IFC’s missions is to promote sustainable private sector investment in developing countries, helping to reduce poverty and improve people’s lives. The types of financial investments provided are loans and debt securities, equity investments and guarantees. The total portfolio invested in MFIs is \$225,000,000 (as of September 2004) allocated in 52 institutions.

### *5.2. Are MFIs Self-Sustainable?*

Sustainability is generally considered in two levels. The first is operational sustainability, which refers to the ability of institutions to generate enough revenue to cover operating costs—but not necessarily the full cost of capital. If unable to do this, capital holdings are depleted over time. The second level of concern is financial sustainability, defined by whether or not the institution requires subsidized inputs in order to operate. The three main pillars that sustainability rests on are:

1. Voluntary deposits.
2. Rational pricing policy, with interest rates on loans high enough to cover costs yet low enough for micro enterprises to afford. Interest rates on deposits should exceed the return those resources would earn if saved in other ways, while interest rates on loans should undercut those of informal moneylenders but exceed those of commercial banks.
3. High repayment rates. This requires, in turn, the promotion of financial discipline among borrowers. The organization must restrain costs and stress efficiency in assessing investment plans, screening borrowers, processing loans, collecting repayments and mobilizing savings.

Nevertheless, it is important to bear in mind that virtually all microfinance institutions in the world charge positive real interest rates, not

subsidizing the borrower (Schreiner and León, 2002). In spite of this practice, some microfinance institutions still have difficulties matching their financial and operational costs due to the high transaction costs that characterize the microfinance technology and/or lack of scale.

One of the main references to the sustainability level of the microfinance institutions is the Microbanking Bulletin, a semi-annual publication dedicated to the financial performance of organizations that provide banking services for the poor (The Microbanking Bulletin, 2003). It is important to state that the data used in the bulletin is self-reported and that their sample of organizations has a strong bias toward those institutions trying to achieve financial sustainability. This publication classifies the MFIs according to three operational criteria: region, scale and target markets.

First, it is important to notice the relatively low default rate of most of these institutions (Table 11-3, Column 1), which ranges from 0.2% to 2.6%. It is noteworthy that in virtually all regions the institutions with the lowest default rates within their region were the smallest institutions which focused on the low end of the market.

In terms of operational and financial sustainability, the figures provide more mixed results. While most of the institutional categories are operationally self-sufficient (OSS), the same cannot be said about their financial self-sufficiency (FSS). Among the institutions that were not operationally sustainable were the African small MFIs and the Latin American small institutions which focused on the low end of the market with 85% and 79% of their operational costs covered (Table 11-3, Column 3). The main group of institutions which presented operational sustainability, the large MFIs, also maintained their performance in terms of financial sustainability, covering 115–133% of their financial costs (Table 11-3, Column 4).

It is also interesting to note that the institutions with a particular focus on small businesses presented a much higher average loan size, \$2,719.00, and were both operationally and financially self-sufficient with percentages of 128 and 108, respectively.

### *5.3. What are the MFI Hurdles to Reach Sustainability?*

Suppliers of microfinance for micro enterprises face two hurdles to reduce costs to reach self-sustainability. First, the institution cannot inexpensively evaluate a borrower's credit risk. Second, micro enterprise borrowers cannot provide the iron-clad guarantees that would reduce the need to evaluate risk.

To ameliorate the problems of imperfect information and weak guarantees, microfinance has resorted to providing homogeneous products. The goal of product design has not been to adjust to the heterogeneity of demand but rather

TABLE 11-3 *Performance indicators of selected microfinance institutions by region*

| MFIs by region, scale of operation and target market |          | Portfolio at risk > 90 days (%) | Average loan balance per borrower (US \$) | Operational self-sufficiency (OSS) (%) | Financial self-sufficiency (FSS) (%) |
|--|----------|---------------------------------|---|--|--------------------------------------|
| All MFIs ( <i>n</i> = 124)                           |          | 1.5                             | 532                                       | 115                                    | 104                                  |
| FSS MFIs ( <i>n</i> = 66)                            |          | 1.5                             | 621                                       | 140*                                   | 128*                                 |
| 1. Africa large                                      | <i>n</i> | 4                               | 6   | 6                                      | 6                                    |
|  | avg      | 1.4                             | 423                                       | 148                                    | 133                                  |
|  | std      | 0.6                             | 307                                       | 33                                     | 30                                   |
| 2. Africa medium                                     | <i>n</i> | 8                               | 8   | 8                                      | 8                                    |
|  | avg      | 1.4                             | 191                                       | 103                                    | 95                                   |
|  | std      | 1.4                             | 140                                       | 21                                     | 22                                   |
| 3. Africa small                                      | <i>n</i> | 5                               | 7   | 7                                      | 7                                    |
|  | avg      | 1                               | 102                                       | 85                                     | 78                                   |
|  | std      | 0.4                             | 54  | 21                                     | 22                                   |
| 4. Asia large  | <i>n</i> | 4                               | 4   | 4                                      | 4                                    |
|  | avg      | 2.2                             | 394                                       | 147                                    | 130                                  |
|  | std      | 2.7                             | 274                                       | 45                                     | 36                                   |
| 5. Asia medium                                       | <i>n</i> | 5                               | 7   | 7                                      | 7                                    |
|  | avg      | 1.6                             | 97  | 111                                    | 100                                  |
|  | std      | 1.7                             | 47  | 10                                     | 18                                   |
| 6. Asia small broad                                  | <i>n</i> | 3                               | 5   | 5                                      | 5                                    |
|  | avg      | 2.6                             | 340                                       | 148                                    | 135                                  |
|  | std      | 2.4                             | 118                                       | 39                                     | 32                                   |
| 7. Asia small low-end                                | <i>n</i> | 3                               | 5   | 3                                      | 5                                    |
|  | avg      | 0.2                             | 75  | 106                                    | 104                                  |
|  | std      | 0.3                             | 43  | 62                                     | 81                                   |
| 8. ECA large   | <i>n</i> | 3                               | 5*  | 5                                      | 5                                    |
|  | avg      | 0.4                             | 1584*                                     | 134                                    | 115                                  |
|  | std      | 0.6                             | 419*                                      | 23                                     | 27                                   |
| 9. ECA medium  | <i>n</i> | 7*                              | 7   | 9                                      | 9                                    |
|  | avg      | 0.2*                            | 739                                       | 121                                    | 110                                  |
|  | std      | 0.2*                            | 400                                       | 31                                     | 38                                   |
| 10. ECA small  | <i>n</i> | 4                               | 6   | 6                                      | 6                                    |
|  | avg      | 0.3                             | 354                                       | 110                                    | 93                                   |
|  | std      | 0.5                             | 198                                       | 35                                     | 32                                   |
| 11. LA credit unions                                 | <i>n</i> | 9                               | 11*                                       | 11                                     | 11                                   |
|  | avg      | 2.1                             | 1898*                                     | 108                                    | 96                                   |
|  | std      | 1.5                             | 875*                                      | 10                                     | 8                                    |
| 12. LA large   | <i>n</i> | 12                              | 2*  | 12                                     | 12                                   |
|  | avg      | 2.6                             | 959*                                      | 129                                    | 127                                  |
|  | std      | 1.8                             | 480*                                      | 19                                     | 21                                   |
| 13. LA medium  | <i>n</i> | 6                               | 8   | 8                                      | 8                                    |
|  | avg      | 1.6                             | 385                                       | 128                                    | 123                                  |
|  | std      | 1.4                             | 240                                       | 28                                     | 28                                   |

TABLE 11-3 *Continued*

| MFIs by region, scale of operation and target market |          | Portfolio at risk > 90 days (%) | Average loan balance per borrower (US \$) | Operational self-sufficiency (OSS) (%) | Financial self-sufficiency (FSS) (%) |
|--|----------|---------------------------------|---|--|--------------------------------------|
| 14. LA small broad                                   | <i>n</i> | 6                               | 6   | 8                                      | 8                                    |
|  | avg      | 2.1                             | 606                                       | 103                                    | 90                                   |
|  | std      | 0.8                             | 461                                       | 27                                     | 24                                   |
| 15. LA small low-end                                 | <i>n</i> | 8                               | 10  | 10*                                    | 10*                                  |
|  | avg      | 1.6                             | 267                                       | 79*                                    | 72*                                  |
|  | std      | 2.4                             | 205                                       | 29*                                    | 28*                                  |
| 16. MENA   | <i>n</i> | 7                               | 7   | 9                                      | 9                                    |
|  | avg      | 0.5                             | 286                                       | 113                                    | 101                                  |
|  | std      | 0.8                             | 154                                       | 36                                     | 31                                   |
| 17. WW small business                                | <i>n</i> | 4                               | 4*  | 4                                      | 4                                    |
|  | avg      | 1.6                             | 2719*                                     | 128                                    | 108                                  |
|  | std      | 2.2                             | 922*                                      | 28                                     | 33                                   |

Region: Africa, Asia, Eastern Europe and Central Asia (ECA), Latin America (LA), and Middle East and North Africa (MENA), and the WW for MFIs that operate in more than one region.

Target market: MFIs are classified into four categories: low-end, broad, high-end and small business—according to the range of clients they serve based on average outstanding loan size in relation to GNP per capita (i.e., depth).

Scale of operations: MFIs are classified as small, medium and large according to the size of their loan portfolio within their regional context to facilitate comparisons of institutions with similar outreach.

For “All MFIs,” averages are calculated on the basis of the values between the second and ninth deciles. Abbreviations: *n* = number of cases; avg = average; std = standard deviation. For “FSS” and each peer group, averages are calculated by dropping the top and bottom observations. Significant differences from the average at the 1% significance level are marked with an asterisk (\*). Additional statistical information is available at [www.microbanking-mbb.org](http://www.microbanking-mbb.org). For further definitions of terms, refer to *The Microbanking Bulletin* (2003, p. 61).

to solve the serious information problems entailed in lending to businesses of this type (Lepp, 1996).

Rhine and Rotblatt (1994) suggested the presence of several common features on successful microfinance organizations, including BancoSol in Bolivia, Actuar/Bogotá in Colombia, Grameen Bank in Bangladesh and the *unit-desa* system in Indonesia.

Invariably, these institutions have positive (and high) real interest rates for both loans and deposits. Careful screening of loan applicants and strict repayment enforcement shows concern for incentives. Guarantee requirements recognize that traditional collateral is rarely available; instead, character references, loan officer judgments of entrepreneurial ability and/or joint-liability mechanisms are used. Financial incentives encouraging prompt repayment

include rebates for timeliness, penalties for late payments and collateral substitutes such as compulsory savings.

Successful intermediaries tend to decentralize decisions while maintaining clear connections between retail outlets and the central office. Decentralized decisions speed processing and place accountability on those employees that are closer to the clients. Retail outlets tend to be small, simple and standardized, and are easily replicable and placed close to clients in economic activity centers. Clients develop personal relationships with loan officers which help them to create personal identification links to lender thus softening the impersonality of otherwise standardized operations. On their side, the central office provides financial management, oversight and technical support.

## 6. DO THEY MAKE AN IMPACT?

Analysis of the impact of MFIs is useful for two main reasons. First, it allows policy makers and practitioners to determine what effects these institutions actually have on their intended outcomes and whether there are important unintended effects. Second, good impact assessment goes beyond pointing out the intended and unintended effects, but must also propose plausible explanations for them. This is crucial for the improvement of the effectiveness of these institutions as it provides important information for the redesign of the programs and products offered.

Most of the impact literature on microfinance has been donor driven and can be placed within a much broader debate about the role and value of development efforts. Aid agencies in particular are under pressure from their executive boards to demonstrate effectiveness and to demonstrate on-the-ground impact on the lives of the people in developing countries (UNDP, 2001; United Nations, 2002). As the development community focuses on issues of relevance and impact, it becomes clear that taxpayers in the developed world and their representatives are no longer willing to support development efforts without a more robust demonstration of its value (Meltzer Commission, 2000). It is unclear, however, whether the impact assessment literature has yet permeated to the donors.

This section briefly discusses some issues in the measurement of impact evaluation of microfinance interventions. This is followed by evidence drawn from the empirical literature. It then presents a summary of the depth of outreach and describes some particularities of impact in developed countries.

### 6.1. How to Measure Impact?

A quantitative *impact evaluation* assesses the changes in the well-being of individuals that can be attributed to a particular program or policy based



on information collected on program participation. It is aimed at providing feedback and helping improve the effectiveness of programs and policies (Baker, 1999). Impact evaluations are decision-making tools for policymakers that enable programs to be made accountable to the public. Such a causal analysis is essential for understanding the relative role of alternative program interventions in reducing poverty. As Hulme (2000) argued:

*Behind all microfinance programs, and indeed virtually all aid financed initiatives, is the assumption that intervention will change human behaviors and practices in ways that lead to the achievement (or raise the probability of achievement) of desired outcomes (Hulme, 2000, p. 81).*

Evaluating the impact of a policy or program hinges on asking the fundamental question: What would have been the situation if the intervention had not taken place? Although one obviously cannot observe such a situation, it is possible to approximate it by constructing an appropriate counterfactual. A counterfactual is a hypothetical situation that tries to depict the welfare levels of individuals in the absence of a policy or program. How a counterfactual is constructed or visualized depends on a number of factors, including program coverage.

For most partial coverage programs, counterfactuals are simulated by comparing program participants (the *treatment group*) with a *control* or *comparison group*. The control or comparison group is made up of individuals (or other units of analysis, such as households, or communities) that do not participate in the program being evaluated but have the same characteristics of the program beneficiaries, especially with respect to those characteristics that are relevant to program participation and program outcomes.

As Baker (1999) observes, the key issue when evaluating the impact of partial-coverage programs is how to select or identify nonparticipants. The group can either be selected randomly through a process similar to a lottery or constructed using special statistical techniques. The choice of the method to identify the group of nonparticipants determines the *evaluation design*, which can be broadly classified into three categories: experimental, quasi-experimental and nonexperimental. These evaluation designs vary in feasibility, cost and the degree of clarity and validity of results. One of the key aspects of all these techniques is how well they address the problem of selection bias. In impact evaluations this is caused by the fact that program participants differ from non-participants in characteristics that cannot be observed by the evaluator and affect both the decision to participate in the program and its outcome (e.g., ability or motivation) (Mosley, 1997).

Selection bias may also go in the opposite direction, however. Many microfinance institutions target women and poor households. At the extreme,

effectively targeting poor households might yield the impression that participation in the program makes the clients poorer (Morduch, 1999).

Another important source of bias is nonrandom program placement (also known as location bias). Many programs are set up specifically to serve the under-served (Coleman, 1999). This may lead to apparent negative impact relative to control areas. Alternatively, the program may set up where good complementary infrastructure is available, biasing estimates upward. The sign and size of the biases are likely to change as programs expand into new areas over time (Morduch, 1999).

The problem of selection bias arises because program participants may be individuals who have the most to gain from a particular program and are more motivated to commit themselves to program activities. This bias can be rather large and several authors show in microfinance literature that not controlling for self-selection can lead to overestimation of the effect of participation by as much as 100% (Pitt and Khandker, 1998; Alexander, 2001; McKernan, 2002).

Karlan (2001) highlights the perils of using quasi- and nonexperimental evaluation designs to assess the impact of microfinance institutions when only cross-sectional data is available. Selection and institutional dynamics problems are difficult to deal with, depending on the circumstances of a given project and the economic setting. One important implication is that, rather often, findings cannot be easily attributed to the project.

## 6.2. *What is their Impact?*

Below some of the main findings from the impact literature are discussed. The first main cluster of studies focuses on the Bolivian experience. Mosley's (1996, 2001) studies have arguably been the ones with the greatest focus on the effect of microfinance on poverty. Not only does this author calculate the poverty level of the clients of the institutions, but he also estimates the effect of the loan on the individuals' enterprises and households. However, his extremely limited datasets—around 24 and 45 observations per institution—and poor control groups substantially restrict the generality of his arguments (Morduch, 1999, p. 1600).

Among his many findings, Mosely (2001) argued that the microfinance institutions Banco Sol, Pro Mujer, PRODEM and Sartawi, have increased their poor clients' income in 51%, 28%, 36% and 13%, respectively, compared with the respective control group. In terms of assets change, the impact was even more substantial, 124%, 86%, 71% and 46%. However, as already mentioned, Mosley's sample is too limited from which to generalize.

Vogelgesang (2001) analyzed the impact of the loans from the microfinance organization Caja Los Andes on its client's enterprises. The study, based on the client's records from this organization, found that on average loan

renewal increased total sales from 3.5% to a statistically significant 15%. While the study estimators were designed to correct for bias arising from the bank's decision to approve a loan application and from the client's decision to apply for repeat loans, the work cannot correct for selection bias arising from the client's decision to apply for a loan since only applicants are observed. Thus, when interpreting the results the author restricts the analysis to micro entrepreneurs willing to take a loan.

Another cluster of studies focuses on Guatemala (Wydick, 1999a, 1999b; Kevane and Wydick, 2001). Data were taken from 260 borrowers from FUNDAP, an ACCION-affiliated credit program in Guatemala and those of 82 entrepreneurs displaying similar characteristics to FUNDAP borrowers but located in areas just outside the reach of FUNDAP's credit program. The studies in this cluster analyze the theoretical and empirical impact of microenterprise credit on child schooling, occupational mobility and the trade-off between growth and poverty alleviation among female entrepreneurs.

Although the authors do not provide any hard evidence, they seem to believe that the micro entrepreneurs served by FUNDAP are poor. In general, the studies argue that credit had a positive and statistically significant effect on the micro entrepreneurs, although some qualification should be made. The authors are aware of the limitation and caveats of this type of analysis and address issues such as selection bias as they argue that their treatment and control groups are not different in terms of their socio-economic characteristics.

Empirical results show that individuals who received credit had a greater probability of moving to self-employment to become an employer than those that did not take credit. The results also suggest that the combined effect of innate entrepreneurial ability and credit access have a greater impact on upward class structure mobility than the interaction between formal schooling and credit access. Moreover, each additional year of credit increased the household enterprise income in 34.1% (Wydick, 1999a).

Keave and Wydick's (2001) research shows that credit had both a positive and statistically significant impact on both employment generation and sales of 0.809 and 1.39, respectively. The answer to their second research question—whether targeting micro enterprise credit at women involves sacrificing economic growth in favor of poverty alleviation and the welfare of children—was that in certain circumstances, this trade-off does exist. In their twenties, male entrepreneurs are predicted to add between 0.2 and 0.35 more employees than female entrepreneurs of the same age, while female entrepreneurs between ages 45–60 are predicted to add about 0.35 to 0.5 more employees to their enterprises than their male counterparts. Nevertheless, the authors qualify that it is within this subset of women that increases in income are likely to have the greatest positive effect on the welfare of children.

Aroca, Andreassi and Romani (2002) measured the impact of nine microfinance programs in Chile and Brazil (i.e. CEAPE, Social-Cred, Banco do Povo-St. André, Bancris, Micro-Cred, Banco de Desarrollo, PROPESA, Banco Estado, Banefe). The authors ran a survey on the clients of these institutions and used this information to match them to self-employed individuals from the Chilean and Brazilian household surveys (i.e., CASEN and PNAD) using the propensity score matching technique (Rosenbaum and Rubin, 1983). The impact evaluation suggested that such programs have a substantially positive and significant effect on their clients in Brazil, generating an average increase in business profits from R\$971 to R\$1,931. In Chile, they reported a 38% increase in the average income of the micro entrepreneurs and a 50% drop in the income of the clients of NGOs.

Azevedo (2002) evaluated the impact of microfinance organizations in Brazil. Using the clients' records of two institutions and building two different comparison groups (reflexive comparison and propensity score matching), the study showed that both institutions had positive and statistically significant effects on their clients' enterprise profits. These averaged R\$748 and R\$1,699 respectively, when using household survey data as a control group, and R\$218 and R\$599 when using a reflexive comparison group.

Another cluster of studies focused on the Peruvian microfinance institution Mi Banco (Alexander, 2001). The data used to analyze this institution was collected as part of the AIMS project, sponsored by USAID (Dunn and Arbuckle Jr., 2001) and consisted of a panel of clients and a control group between the years 1997 and 1999.

Alexander (2001) reinforced the previous findings of Coleman (1999) and Pitt and Khandker (1998) who observed that estimates which ignore selection bias will overstate the impact of credit. After controlling for selection bias, however, Alexander finds that credit leads to higher levels of micro enterprise profits for those micro entrepreneurs who choose to access Mi Banco loans. These show robust results across both weekly and monthly enterprise profits. The author fully explores the possibility of this dataset and estimates a household fixed effects model, which was previously unavailable for large impact assessments in this literature. This estimate strengthens the earlier findings by showing that benefits of microfinance loans are large and significant.

Copestake, Bhalotra and Johnson (2001) estimate the impact of an urban credit program in Zambia on business performance and on a range of well-being indicators. Borrowers who obtained a second loan experienced significantly higher average growth in business profits and household income. Inflexible group enforcement of loan obligations resulted in some borrowers, especially those who had taken only one loan, being made worse off. Their methodological investigations suggest that the supply of rigorous impact studies

can be increased by basing them on data collection that serves a wider range of purposes, including market research.

MFC (2000) analyzed the impact of the polish MFI Inicjatywa Mikro (IM). This MFI, which began its operation in March 1996, operates exclusively in southern Poland (Krakow, Katowice and Bielsko Biala). IM's activity focuses on major cities in Eastern Europe. The target population is small enterprises that employ up to ten people and have potential growth but do not qualify for bank credit because of lack of adequate collateral. The current effective annual interest rate ranges from 24% to 29%. A total sample of 150 households was interviewed in August 1999, out of which 120 were first time (69) and repeated (51) clients and 30 eligible nonclients. Using longitudinal data collected through recall questions for the status during the previous year, the study showed increase in household well being (as measured by fixed assets, expenditures and savings) and improvement in the business (demonstrated through income and fixed assets).

There may also be unintended effects. A few authors have recently provided evidence that microcredit might not always be the best way to help the less advantageous. A handful of studies have shown these institutions might not only increase the vulnerability of the household but also promote negative shifts in gender relations within the family (Goetz and Gupta, 1996; Hulme and Mosley, 1996; Rogaly, 1996; Buckley, 1997; Johnson and Rogaly, 1997; Zeller and Sharma, 2000; Pretes, 2002).

### *6.3. How poor are the Clients?*

Another relevant empirical question is how poor the "poor" are who use microfinance institutions. Most of the empirical studies in the world suggest that the microfinance institutions are relatively effective at reaching the poor but not the poorest and that other anti-poverty modalities need to be brought into play when attempting to reach the poorest (Navajas, Schreiner, Meyer, Gonzalez-Vega and Rodriguez-Meza, 2000; Mosley, 2001; Smith, 2002).

While virtually all MFIs have a declared objective of providing financial services to the poor, many end up serving the better-off among the poor rather than the very poor. This has led to an increasing interest in methodologies and products that would serve to deepen the outreach of financial services. The limited evidence that exists on the depth of outreach suggests that strong leadership commitment and the use of targeting tools is important.

In the past few years, several studies were conducted using the CGAP Poverty Assessment Tool to determine the poverty levels of entering MFI clients relative to non-client control groups (CGAP, 2001). The tool involved conducting quantitative surveys using a structured household questionnaire to capture key indicators of poverty. Principal Component Analysis (PCA) is then

used to produce composite poverty indexes. These are then used to compare new MFI clients with the control population. The findings indicate that in virtually all cases with an appropriate counterfactual (Bolivia, Peru and Brazil), MFIs served a higher proportion of clients below the poverty line than commercial banks.

Conventional wisdom also suggests that institutional structure matters and that NGO-MFIs are better pre-disposed to work with poorer clients. However, some authors have found that institutional type did not seem to have a determining effect on depth of outreach, while institutional location or particular target groups seemed a lot more influential (Azevedo, Mourji, Ndiaye and Deshpande, 2004).

These results highlight the importance of controlling for community-level variables when analyzing poverty levels. More importantly, they also suggest that it is possible to identify and work with vulnerable clients even in relatively affluent areas. The results also suggest that targeting particular socio-economic groups does have an effect on depth of outreach.

## 7. POLICY IMPLICATIONS

Although MFIs have had a positive impact on households and micro enterprises throughout the world, it is important to remember that the greatest promise of these organizations, namely, their financial sustainability, is still largely unmet. The programs that are breaking even financially are not those celebrated for serving the poorest clients. In other words, the double bottom line, poverty reduction with financial profitability, remains a chimera.

In terms of their clients' poverty level, several institutions presented evidence that a percentage of their clients are below the poverty line. The 2000 household survey of Bolivia shows that 23% of the clients of the Bolivian commercial banks are below the poverty line. In the case of the Fundos Financeiros Privados and Cajas de Ahorros, this figure reaches 40.9% and 39.9%, respectively. This suggests that the technology used by new institutions is making a difference. However, some studies use labor occupation as a proxy for the poverty level of their clients, which is not a direct measure and needs to be improved upon.

The financial vulnerability of these organizations creates a dependence on subsidies given by international and domestic donors, and subsequent pressure from the international community in particular for the sustainability of these institutions. But if the money spent to support microfinance helps to meet social objectives in ways not possible through alternative programs like workfare and direct food aid, why not continue subsidizing microfinance? Could the development effectiveness of microfinance justify subsidies given by the

donor community to these institutions? These are issues that have still not been satisfactorily dealt with in the literature; further empirical evidence is needed.

Answering these questions requires studies of social impact and information on clients' profiles by income and occupation, as well as precise estimates of the level of subsidy demanded by these institutions. For example, we need studies that quantify the average cost of moving one person out of poverty through microfinance institutions compared with alternative categories of anti-poverty expenditure such as primary health, primary education and rural road building.

It seems that those countries that have a greater exposure to international donors are also the ones that have produced the largest number of studies. Domestic donors should also be aware of the importance and necessity of evaluating their policies in order to improve effectiveness and to hit their targets.

Studies also suggest that the provision of complementary services, such as education, health and business development can have a positive effect on the household. One question that remains to be addressed is to what extent nonfinancial activities could be provided by sister institutions instead of by the microfinance organization itself.

Policy makers should be aware that credit is only one among the many different products that microfinance institutions are able to provide to those traditionally excluded from the financial sector. Savings, deposits and money transfers can also play an important role in helping the poor to manage exogenous shocks and thereby decreasing the volatility of their incomes.

To conclude, the available evidence suggests that microfinance institutions do reach the poor. Although in some countries there is still a need for further evidence of how much better microfinance institutions are in comparison with commercial banks in this particular objective, in others (such as Bolivia, Bangladesh and Indonesia) the data suggest that microfinance institutions have a much higher percentage of poor individuals in their portfolio compared with commercial banks. Moreover, these institutions have had a positive impact on their clients' income, assets and employment generation capability. However, it is also known that most of these microfinance institutions are not financially sustainable, requiring loans at concessional rates to operate. Future research needs to determine which extension tie-ins should be internalized, as well as the impact of other services provided by microfinance institutions such as savings and insurance. In addition, future research should explore whether the social benefits of microfinance exceed the social costs.

Microfinance is here to stay, and its place in the toolbox of development policy makers, practitioners, activists and researchers has been consolidated. Nevertheless the limits and potential of this innovative instrument must be continually re-evaluated in order to guarantee its effectiveness.

## NOTES

<sup>1</sup> The performance of Grameen scheme is far from undisputed. One criticism came from an article published in the Wall Street Journal on November 27, 2001, which has been extensively discussed in specialized forums, such as the DevFinance mailing list <http://microfinancegateway.org/website/devfinance/index.htm>.

<sup>2</sup> In 1998, the United Nations General Assembly proclaimed 2005 as the International Year of Microcredit (resolution 53/197 of December 15, 1998).

<sup>3</sup> These are rapidly growing product lines in several developing countries such as Brazil, Bolivia and South Africa, which can have a potentially positive impact on the consumption smoothing out of the poor and/or vulnerable. However, this effect is far from undisputed: see Rhyne (2001) for a thorough analysis of the Bolivian over-indebtedness crisis in the late 1990s.

<sup>4</sup> Orlando and Pollack (2000) showed that only one-fourth of the micro entrepreneurs in Latin America and the Caribbean are deemed to be poor.

<sup>5</sup> According to Westley (2001), 70% of Latin America's poor earners are either employees of microenterprises or self-employed.

<sup>6</sup> CGAP (2004) showed that NGOs, banks and nonbank financial institutions that specialize in microfinance, as well as microfinance programs in full-service commercial banks, represent around 33% of all active loans in these financial institutions.

<sup>7</sup> Actual inward remittances are probably considerably higher because the above figure does not include funds transferred through couriers and other informal mechanisms. In addition, in-country transfer services are also important, especially for rural families supported by a member working in the city (CGAP, 2004, p. 10).

## REFERENCES

- Alexander, G. M. (2001). *An Empirical Analysis of Microfinance: Who Are the Clients?* Mimeo, Washington, DC.
- Aportela, F. (2000). *Effects of Financial Access on Savings by Low-Income People*. Conference Proceedings, Fifth Annual Meeting of LACEA, Rio de Janeiro, Brazil, 12–14 October.
- Aroca, P. A., Andreassi, T. and Romani, G. (2002). *Estudio De Programas De Microcredito En Brazil Y Chile*. Mimeo, University of Illinois. Urbana, IL.
- Azevedo, J. P. (2002). *The Impact of Microfinance Loans on the Clients' Enterprises: A Quasi-Experiment with Clients and Household Survey Data*. Mimeo, Newcastle-upon-Tyne, England.
- Azevedo, J. P. (2003). *The Nature and Performance of Micro-Entrepreneurs: Evidence from Microcredit Clients in the Slums of Rio*. Conference Proceedings, Eighth Annual Meeting of LACEA, Puebla, Mexico. 8–11 October.
- Azevedo, J. P., Mourji, W. F., Ndiaye, A. and Deshpande, G. R. (2004). *Determining the Depth of Outreach of Senegalese MFIs*. CGAP Report, Washington, DC.
- Baker, J. (1999). *Evaluating the Poverty Impact of Projects: A Handbook for Practitioners*. World Bank, Washington, DC.
- Becker, G. S. (1991). *A Treatise on the Family*. Harvard University Press, Cambridge MA.
- Besley, T. and Coate, S. (1995). Group Lending, Repayment Incentives and Social Collateral. *Journal of Development Economics* 46: 1–18.
- Buckley, G. (1997). Microfinance in Africa: Is It Either the Problem or the Solution? *World Development*: 25: 1081–1093.



- Cavalcanti, T. (2000). A Experiência Do Vivacred: Integração Social E Distribuição De Renda. *O Mercado de Trabalho do Rio de Janeiro: Conjuntura e Análise* 8: 5–7.
- CGAP (1997). *How CGAP Member Donors Fund Microfinance Institutions*. Focus Note No. 11, Washington, DC.
- CGAP (2001). *Assessing the Relative Poverty of Microfinance Clients: A CGAP Operational Tool*. Report, Washington, DC.
- CGAP (2004). *Financial Institutions with a “Double Bottom Line”: Implications for the Future of Microfinance*. Occasional Paper, CGAP, Washington, DC.
- CGCED (2002). *Caribbean Economics Overview 2002: Macroeconomic Volatility, Household Vulnerability and Institutional and Policy Responses*. World Bank, Washington, DC.
- Chami, R., Fullenkamp, C. and Jahjah, S. (2005). Are Immigrant Remittance Flows a Source of Capital for Development? *IMF Staff Papers* 52: 55–81.
- Coleman, B. E. (1999). The Impact of Group Lending in Northeast Thailand. *Journal of Development Economics* 60: 105–141.
- Copestake, J., Bhalotra, S. and Johnson, S. (2001). Assessing the Impact of Microcredit: A Zambian Case Study. *Journal of Development Studies* 37: 81–100.
- Daley-Harris, S. (2003). *Pathways out of Poverty: Innovations in Microfinance for the Poorest Families*. Kumarian Press, Bloomfield, CT.
- Dunn, E. and Arbuckle, J. G. Jr. (2001). *The Impacts of Microcredit: A Case Study from Peru*. AIMS-USAID Report, Washington, DC.
- Edgcomb, E. (2002). *Improving Micro enterprise Training and Technical Assistance: Findings for Program Managers*. Micro enterprise Fund for Innovation, Effectiveness, Learning, and Dissemination. Aspen Institute.
- Ghatak, M. (1999). Group Lending Contracts and the Peer Selection Effect. *Journal of Development Economics* 60: 27–50.
- Ghatak, M. (2000). Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect. *Economic Journal* 110: 601–631.
- Ghatak, M. and Guinnane, T. W. (1999). The Economics of Lending with Joint Liability: Theory and Practice. *Journal of Development Economics* 60: 195–228.
- Gill, I., Yusuf, S., Von Amsberg, J., Dillinger, W., Menezes, B. and Neri, M. (1999). *Rio De Janeiro: A City Study*. World Bank, Washington, DC.
- Goetz, A. M. and Gupta, R. S. (1996). Who Takes the Credit? Gender, Power and Control over Loan Use in Rural Credit Programs in Bangladesh. *World Development* 24: 45–63.
- Gulli, H. (1998). *Microfinance and Poverty: Questioning the Conventional Wisdom*. Inter-American Development Bank Sustainable Development Dept. Micro enterprise Unit, Washington, DC.
- Heckman, J. J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica* 46: 931–959.
- Heckman, J. J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *Journal of Human Resources* 32: 441–462.
- Hickson, R. (2001). Financial Services for the Very Poor—Thinking Outside the Box. *Small Enterprise Development* 12: 55–67.
- Hoff, K. and Stiglitz, J. E. (1990). Imperfect Information and Rural Credit Markets: Puzzles and Policy Perspectives. *World Bank Economics Review* 4: 235–251.
- Hulme, D. (2000). Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice. *World Development* 28: 79–98.
- Hulme, D. and Mosley, P. (1996). *Finance against Poverty*. Routledge, London.
- ILO (1999). *The Micro enterprise in Central America*. International Labor Organization, San Jose, Costa Rica.

- IPEA, PNUD and PMRJ (2001). *Relatório De Desenvolvimento Humano Sustentável Na Cidade Do Rio De Janeiro*. IPEA, Rio de Janeiro.
- Jain, S. (1999). Symbiosis Versus Crowding-Out: The Interaction of Formal and Informal Credit Markets in Developing Countries. *Journal of Development Economics* 59: 419–444.
- Johnson, S. and Rogaly, B. (1997). *Microfinance and Poverty Reduction*. Oxford, U.K.
- Kapur, D. (2004). *Remittances: The New Development Mantra?* G-24 Discussion Paper No. 29, UNCTAD, Geneva.
- Karlan, D. S. (2001). Microfinance Impact Assessments: The Perils of Using New Members as a Control Group. *Journal of Microfinance* 3: 75–85.
- Kevane, M. and Wydick, B. (2001). Micro enterprise Lending to Female Entrepreneurs: Sacrificing Economic Growth for Poverty Alleviation? *World Development* 29: 1225–1236.
- Krahnen, J. P. and Schmidt, R. H. (1994). *Development Finance as Institution Building*. Westview Press, Boulder.
- Ledgerwood, J. (1999). *Microfinance Handbook: An Institutional and Financial Perspective*. World Bank, Washington, DC.
- Lensink, R., Gangopadhyay, S. and Ghatak, M. (2005). On Joint Liability Lending and the Peer Selection Effect. *Economic Journal*, forthcoming.
- Lensink, R., Hermes, N. and Mehrteab, H. (2005). Peer Monitoring, Social Ties and Moral Hazard in Group Lending Programs: Evidence from Eritrea. *World Development* 33: 149–169.
- Lensink, R. and Mehrteab, H. (2003). *Risk Behavior and Group Formation in Microcredit Groups in Eritrea*. Research Paper No. 03/02, CREDIT.
- Lepp, A. (1996). Financial Products for Mess—the Municipal Savings and Loan Banks of Peru. *Small Enterprise Development* 7: 15–25.
- McKernan, S.-M. (2002). The Impact of Microcredit Programs on Self-Employment Profits: Do Noncredit Program Aspects Matter? *The Review of Economics and Statistics* 84: 93–115.
- Meltzer Commission (2000). *The Meltzer Commission Final Report*. US Congress, Washington, DC.
- MFC (2000). *Aims and MFC Impact Assessment Project: Tools Testing in the Region of Central and Eastern Europe and the New Independent States*. Microfinance Center for CEEE and the NIS, Warsaw.
- MIF (2002). *Remittances to Latin America and the Caribbean*. MIF/IDB, Washington, DC.
- Morduch, J. (1999). The Microfinance Promise. *Journal of Economic Literature* 37: 1569–1614.
- Mosley, P. (1996). Metamorphosis from NGO to Commercial Bank: The Case of Bancosol in Bolivia. In *Finance Against Poverty*, D. Hulme and P. Mosley (Eds), Routledge, London.
- Mosley, P. (1997). *The Use of Control Groups in Impact Assessments for Microfinance*. Social Finance Unit Working Paper N/19. International Labour Office, Geneva.
- Mosley, P. (2001). Microfinance and Poverty in Bolivia. *Journal of Development Studies* 37: 101–132.
- Mosley, P. and Hulme, D. (1998). Micro enterprise Finance: Is There a Conflict between Growth and Poverty Alleviation? *World Development* 26: 783–790.
- Navajas, S., Schreiner, M., Meyer, R. L., Gonzalez-Vega, C. and Rodriguez-Meza, J. (2000). Microcredit and the Poorest of the Poor: Theory and Evidence from Bolivia. *World Development* 28: 333–346.
- Orlando, M. B. and Pollack, M. (2000). *Micro enterprise and Poverty: Evidence from Latin America*. Inter-American Development Bank: 45, Washington, DC.
- Paetkau, M. (1999). *Financial Support for Micro and Small Enterprises (MSEs) in Thailand*. Micro And Small Enterprise Development & Poverty Alleviation In Thailand—Working Paper 4, International Labor Organization, Geneva.
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge University Press, Cambridge.

- Pitt, M. M. and Khandker, S. R. (1998). The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy* 106: 958–996.
- Pretes, M. (2002). Microequity and Microfinance. *World Development* 30: 1341–1353.
- Rhyne, E. (2001). *Mainstreaming Microfinance: How Lending to the Poor Began, Grew and Came of Age in Bolivia*. Kumarian Press, Bloomfield, CT.
- Rhyne, E. and Rotblatt, L. (1994). *What Makes Them Tick? Exploring the Anatomy of Major Micro enterprise Finance Organizations*. ACCION International Monograph Series No. 9, Washington, DC.
- Rogaly, B. (1996). Microfinance Evangelism, “Destitute Women” and the Hard Selling of a New Anti-Poverty Formula. *Development in Practice* 6: 100–112.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70: 41–55.
- Rutherford, S. (2001). *The Poor and Their Money*. Oxford University Press, Oxford.
- Sadoulet, L. and Carpenter, S. B. (2001). *Endogenous Matching and Risk Heterogeneity: Evidence on Microcredit Group Formation in Guatemala*. Manuscript, ECARES/Free University of Brussels.
- Schreiner, M. (2004a). *Asset-Building for Micro enterprise through Matched Savings in Low-Income Countries*. Paper presented to Symposium on Micro enterprise Solutions for the World’s Poorest, Metropolitan College, New York City.
- Schreiner, M. (2004b). Micro enterprise in the First and Third Worlds. *World Development* 31: 1580–1597.
- Schreiner, M. and León, J. (2002). Microfinance for Micro enterprise: A Source Book for Donors. *Savings and Development* 26.
- Smith, S. C. (2002). Village Banking and Maternal and Child Health: Evidence from Ecuador and Honduras. *World Development* 30: 707–723.
- Stephens, B. (2005). Sustainability in Sight: An Analysis of MFIs That Become Sustainable. *Microbanking Bulletin* 10: 23–29.
- Stiglitz, J. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review* 71: 393–410.
- The Microbanking Bulletin (2003). *Focus on Savings, Issue No. 9*. Microfinance Program Economics Institute, Boulder.
- The Microcredit Summit Report (1997). *The Microcredit Summit Report*. Results Educational Fund, Washington, DC.
- UNDP (2001). *Development Effectiveness: Review of Evaluative Evidence*. United Nations Development Program, New York.
- United Nations (2002). *Report of the International Conference of Financing for Development*. United Nations General Secretary, New York.
- Vogelgesang, U. (2001). *The Impact of Microfinance Loans on the Clients’ Enterprise: Caja Los Andes, Bolivia*. Gk working paper series no. 2001-03, University of Mannheim.
- Von Pischke, J. D. (1991). *Finance at the Frontier: Debt Capacity and the Role of Credit in the Private Economy*. World Bank, Washington, DC.
- Waterfield, C. (1994). *Designing Financial Products for the Mse Sector*. IADB/FUNDES Micro & Small Enterprise Donor Seminar, Inter-American Development Bank, Washington, DC.
- Westley, G. D. (2001). *Can Financial Market Policies Reduce Income Inequality?* Inter-American Development Bank Report, Washington, DC.
- Woodruff, C. and Zenteno, R. (2001). *Remittances and Micro enterprise in Mexico*. Working Paper No. 01-25, Santa Cruz Center for International Economics, Santa Cruz, CA.
- World Bank (2003). *Global Development Finance*. World Bank, Washington, DC.

- Wydick, B. (1999a). Credit Access, Human Capital and Class Structure Mobility. *Journal of Development Studies* 35: 131–152.
- Wydick, B. (1999b). The Effect of Micro enterprise Lending on Child Schooling in Guatemala. *Economic Development and Cultural Change* 47: 853–869.
- Zeller, M. and Sharma, M. (2000). Many Borrow, More Save and All Insure: Implications for Food and Micro-Finance Policy. *Food Policy* 25: 143–167.

## Stage 4: Venture Development I: Private Sector Issues

SIMON C. PARKER

*Durham Business School, Mill Hill Lane, Durham DH1 3LB, United Kingdom*

*E-mail: s.c.parker@durham.ac.uk*

## 12. Entrepreneurs as Producers

### 1. INTRODUCTION

Why are some entrepreneurs successful while others are not? Success—whether measured in terms of market share, profits or some other indicator of performance—is the goal of the aspiring entrepreneur and is commonly held up by the policy community as a sign of economic vitality. To this end, policy makers and academics are keen to understand why some entrepreneurial ventures generate high sales and why others do not; why some employ many workers while others remain small; and which ventures choose to make large capital investment decisions. At heart, all of these issues come down to understanding the behavior of the entrepreneur as a producer.

There has been an unfortunate tendency in some parts of the entrepreneurship research field to dismiss production as an irrelevant and remote sub-area of microeconomics, asserting that there is nothing “entrepreneurial” about merely *running* a venture. This viewpoint is related to criticisms of microeconomics as a discipline which, it is claimed, can predict how optimizing competitive firms should organize their production decisions, but are silent about how entrepreneurs behave (e.g., Kirchoff, 1991).

What is unfortunate about this viewpoint is that, whatever entrepreneurship scholars might think, governments are deeply concerned about how many jobs new entrepreneurial ventures create, how much they invest, and how much wealth they generate in their local and national economies. They are concerned because they commonly believe that entrepreneurship can provide a route out of poverty and social exclusion, generating jobs and prosperity that benefit their local communities. Understanding how entrepreneurs can and should organize their factors of production in a way that promotes success

should therefore be more properly regarded as one of the most fundamental aspects of entrepreneurship research.

Despite this, only a limited amount of research has focused directly on these issues to date. In an attempt to fill this gap, the present chapter will attempt to make the case that a production approach to entrepreneurship encompasses a broad range of issues that bear directly on what entrepreneurs do and how they behave. It also forms an integral component of venture development and has generated numerous interesting empirical findings as well as promising many more in the future.

At the outset, it might be helpful to distinguish between two distinct ways of modeling the entrepreneurial production process: as static and dynamic phenomena. According to a purely static view of entrepreneurship, production technology can be treated as effectively fixed. In contrast, a dynamic perspective recognizes that production technologies can and do change over time, and that the entrepreneur has some discretion over this process of change. Both perspectives offer valuable insights and both will be treated in this chapter.

The chapter starts with the static approach. It describes how entrepreneurs marshal the factors of production available to them to produce goods and services. These factors invariably include the entrepreneur's own effort, and sometimes also capital, hired labor and land. In the section after, I consider the active role of the entrepreneur in the production function. I discuss four constructive ways that the neoclassical production function can be extended. These include discontinuous Schumpeterian change as well as more gradual endogenous innovation; the incorporation of idiosyncratic talent and ability; and the introduction of "entrepreneurship capital," a concept recently proposed by Audretsch and Keilbach (2004). I will argue that all of this work implies that the neoclassical economics framework, suitably extended, can encompass many different aspects of entrepreneurship, not excluding the Schumpeterian one of radical change in production technology. Perhaps most controversially, I claim that the neoclassical approach is not rendered redundant by Schumpeterian innovation. Rather, this approach is capable of representing entrepreneurs as idiosyncratic decision-makers instead of passive optimizing automatons, as has been claimed in the past (Baumol, 1968).

After treating the nature of entrepreneurs' production functions as just described, the chapter focuses on one factor of production in particular: entrepreneurs' own labor input. Unlike the others, this particular input is always present in all enterprises. I discuss current and topical issues in entrepreneurs' work effort, both theoretically and with reference to recent empirical findings. The penultimate section briefly reviews evidence that connects production and entrepreneurial performance. The final section concludes the chapter.

## 2. THE NEOCLASSICAL PRODUCTION FUNCTION

Consider the following production relationship:

$$y = f(K, L_o, L_e, M), \quad (1)$$

where  $y$  is output produced in a given period,  $K$  is a capital input,  $L_o$  is the input of the entrepreneur's own labor into the enterprise,  $L_e$  is the input of external, hired, labor in the period, and  $M$  is land. The function  $f$  relates the factor inputs to output and is called the "production function."

Standard assumptions about the production function are that it is increasing in all its inputs, but that *diminishing marginal returns* to those inputs eventually set in. That is, if every factor except one were held fixed, then increasing the amount of this one factor would increase output by progressively smaller amounts as more of the factor was applied. Further assumptions about the technical structure of  $f$  are commonly made in economics, such as homotheticity, or homogeneity with respect to factor inputs (see, e.g., Varian, 1992). None of these assumptions will be discussed any further in this chapter.

Under the above assumptions, it is possible to deduce the amounts of each factor that an entrepreneur will choose to maximize expected profits. So, for example, if  $p$  is the output price, and  $w$  is the price of hired labor (the "wage" rate), then the first-order condition

$$p \partial f / \partial L_e - w = 0 \quad (2)$$

determines the optimal demand for labor given an elastic supply of this factor (Hamermesh, 1993). Demand for the other inputs is determined in the same way. So in principle, once we know the production technology  $f$ , we can predict how much labor and capital a given entrepreneur should hire at given factor prices.

Strictly speaking, the technical relationship  $f$  describes the *maximum* output that an entrepreneur can obtain from her inputs, that is, assuming that he/she operates efficiently. If she did not, then the equality sign in (1) would be replaced by the less than sign. In competitive markets, only the efficient survive so it seems reasonable to suppose that in the long run the equality in (1) is appropriate. In practice, however, government regulations that impede efficient resource usage (e.g., maximum working weeks and employment protection legislation) might prevent *all* ventures from operating at maximum efficiency. This might necessitate a modification to (1) of the form

$$y = f(K, L_o, L_e, M) - u, \quad (1')$$

where  $u \geq 0$ , reflecting deviation of actual production from the maximum (see, e.g., Aigner et al., 1977). This type of production *frontier model* is



rarely estimated for entrepreneurial firms (see Alvarez and Crespi, 2003, for an exception).

Equation (1) is a common specification of the production function, although it is both less and more restrictive than is necessary in the context of entrepreneurship. It is less restrictive by including factors of production that are often irrelevant to entrepreneurs. For example, in most developed countries only a minority of entrepreneurs employs any other workers. For most entrepreneurs, therefore,  $L_e$  is redundant. Likewise, for nonagricultural businesses, land is rarely regarded as an important factor of production. So for all intents and purposes,  $M$  is redundant as well. However, (1) is arguably more restrictive than it needs to be for many entrepreneurial ventures, as I go on to discuss in the next section. In particular, (1) ignores the possibility of technological change, innovation, entrepreneurs' choice of  $f$ , entrepreneurial ability and "entrepreneurship capital," which are all potentially important phenomena in the long run analysis of venture development.

In the remainder of this section, I will say a little about each of the factor inputs  $K$ ,  $L_e$  and  $M$ . The motivation is to emphasize some aspects of production relationships for typical entrepreneurs which I take broadly to be the people who run and develop their own ventures.

### 2.1. Capital

According to Meyer (1990), 60% of new entrepreneurs have no verifiable depreciable capital. The remainder use varying amounts of capital inputs which vary with industry sector and occupation. For example, entrepreneurs who operate a consulting business rarely need to employ much capital: a computer and printer, perhaps, and a telephone. In contrast, an entrepreneur who runs a large manufacturing business has to invest in at least enough capital to reach the minimum efficient scale in their industry. By and large, the entrepreneurship literature has paid relatively little attention to the role of physical capital, with three broad exceptions. One relates to capital constraints whereby entrepreneurs with positive-net present value projects are unable to convince lenders to advance them the sums required to purchase fixed and working capital. This topic has been surveyed elsewhere (e.g., Parker, 2004, Chapter 7) and since it is peripheral to the present inquiry it need not detain us here. The second issue that has been discussed in the literature relates to "minimum efficient scale" and industry entry and exit; see Chapter 7 of this volume by Martin Carree. The third relates to the effects of aggregate capital accumulation on the decision to become an entrepreneur and the average size of firms. Lucas (1978) explored this issue using a production function that extends (1) to include entrepreneurial ability. I will describe Lucas's extended production function below. For now, it merely suffices to summarize Lucas'

key prediction. If the *elasticity of technical substitution* between labor and capital is less than one, as evidence suggests (Hamermesh, 1993), then Lucas' model predicts that as economies accumulate capital, small-scale entrepreneurs are progressively squeezed out by larger firms. Large firms deploy additional capital more efficiently, bidding up the wage and encouraging the entrepreneurs with the smallest ventures to quit entrepreneurship and take wage employment instead. According to this story, venture development involves continued capital accumulation by initially large firms, while smaller ventures (voluntarily) leave the market altogether.

## 2.2. *Hired Labor*

Only a minority (20–30%) of self-employed business owners employ any paid workers (see Carroll et al., 2000; Kuhn and Schuetze, 2001; and Moralee, 1998, for evidence from the U.S., Canada, and the U.K., respectively). A higher proportion of the self-employed use hired labor in some of the countries of continental Europe, reaching as high as 46% in Denmark and 51% in Germany (Cowling, 2003). These rates typically exclude the work input of unpaid family workers, including that of spouses. Theoretical and empirical work on the determinants of entrepreneurs' labor demand is very limited, though there are a handful of exceptions (e.g., Jefferson, 1997; Carroll et al., 2000; Cowling et al., 2004). In principle, labor demand can be analyzed in the neoclassical tradition as explained above, deriving a labor demand function from (2) which depends on  $w$  and possibly also several other aspects of the entrepreneur's venture. An interesting question is why some ventures develop along the employment route, and why others do not. At the time of writing, our knowledge of this important phenomenon is rather poorly developed, although according to van Praag and Cramer (2001) and Cowling et al. (2004), "job creators" are more likely to be male, well educated and in possession of prior self-employment experience.

## 2.3. *Land*

This factor of production affects mainly agricultural businesses, such as farms, although to the extent that it also includes natural resources, other types of venture might also fall within its scope. For example, Schumpeter (1934) mentioned "capturing a new source of supply" as one of several vehicles for entrepreneurship. Apart from farms or real estate companies, however, the acquisition of land is not a key aspect of production or venture development for most entrepreneurial firms. True, expansion of a plant might require a new site on which new equipment or an expanded production process is located; but here land is only a secondary factor, that is, it follows the demands for

capital or labor. For their part, agricultural businesses often fall outside the scope of entrepreneurship studies altogether as they differ markedly from most nonagricultural concerns (though see Jacoby, 1993, for an exception).

### 3. INTRODUCING THE ENTREPRENEUR: EXTENDING THE NEOCLASSICAL PRODUCTION FUNCTION

So far, the entrepreneur has been treated as a passive decision-maker who merely decides how many productive inputs to use. But even this might overstate their scope for action. Entrepreneurial ventures are typically too small to influence output or factor prices, so entrepreneurs are often price-takers rather than price-makers. To survive in competitive markets, the levels of productive inputs are therefore effectively decided by the market as well, since if an entrepreneur makes sub-optimal choices, a more efficient rival can undercut them and steal their business. Thus entrepreneurs do not even have to know how to optimize. Trial and error is sufficient to locate the optimal factor inputs that guarantee survival.

The above arguments might explain why some entrepreneurship scholars have criticized the neoclassical approach for sidelining the entrepreneur. Critics argue that the neoclassical approach does not envision the entrepreneur as an innovator, an opportunity-spotter or a disrupter of smooth production relationships. The idiosyncratic flair of the individual entrepreneur, it is claimed, is ignored in this framework. As Baumol put it: “*one hears of no . . . brilliant innovations, of no charisma or any of the other stuff of which entrepreneurship is made*” (1968, p. 67).

In fact, although contemporary entrepreneurship scholars still cite these and similar arguments to attack neoclassical economics, many of these criticisms are out of date and fall wide of the mark. The economics literature has developed rapidly in the last three decades, and as I argue in this section, the neoclassical production function can be extended in several ways to allay these concerns by:

- 3.1. Specifying and explaining discontinuous shifts whereby entrepreneurs adopt completely new production functions as they develop their venture.
- 3.2. Modeling how entrepreneurs innovate and change their technology gradually.
- 3.3. Introducing idiosyncratic entrepreneurial ability into the production function.
- 3.4. Recognizing that productivity in entrepreneurship might be enhanced by “entrepreneurship capital.”

### 3.1. Discontinuous Shifts

To show how the neoclassical approach can be extended to take account of discontinuous shifts in the production function, consider the following simple story. At time  $t$ , an entrepreneur is aware of a set of feasible production technologies, denoted by  $\mathfrak{S}(t)$ . By *feasible*, I mean that the entrepreneur can adopt any of them they wish, although they can only operate one of them at any point in time. The dependence on  $t$  indicates that technological opportunities evolve through time, which can be thought of as reflecting technological progress. Suppose without loss of generality that, at any time  $t$ , the entrepreneur's best guess about the future technology set  $\mathfrak{S}(t + s)$  is  $\mathfrak{S}(t)$ , for all  $s > 0$ .

At the time of start-up, the choice of technology is unconstrained, but once an entrepreneur has put a technology in place, there is a fixed cost  $\phi$  of replacing it with an alternative. Denote the initial chosen production function by  $f_j \in \mathfrak{S}(0)$ . The profit derived from using production function  $f_i \in \mathfrak{S}(t)$ ,  $i = 1, 2, \dots$  at time  $t$  is  $\pi_i(t) > 0$ . If  $r$  is the interest rate facing the entrepreneur, and  $p_m$  is the price of land, then

$$\pi_i = p f_i(K, L_o, L_e, M) - r \cdot K - w \cdot L_e - p_m M.$$

An entrepreneur will switch from technology  $j$  to technology  $i^*$ , where

$$i^* = \operatorname{argmax} \pi_i, \tag{3}$$

if and only if both (a)  $i^* \neq j$  and (b)  $\pi_{i^*} - \pi_j > \phi$ .

Notice that a discontinuous shift in technology can have dramatic consequences. As the production function  $f_i$  changes, so do the optimal factor demands  $K^*$ ,  $L_e^*$ ,  $M^*$ , which are given as the solutions to

$$p \cdot \partial f_i / \partial K - r = p \cdot \partial f_i / \partial L_e - w = p \cdot \partial f_i / \partial M - p_m = 0. \tag{4}$$

Very large shifts in resource use can be entailed if entrepreneurs switch production function  $f_i$ . For example,  $f_j$  might be a labor-intensive technology, whereas  $f_{i^*}$  is capital-intensive. Then switching from the former to the latter technology can lead to mass redundancies.

The story so far has dealt with a single entrepreneur. In practice, not all entrepreneurs are likely to be aware of the feasible technology set,  $\mathfrak{S}(t)$  (Kirzner, 1985). Some (perhaps most) entrepreneurs might be ill informed, having access to a subset  $\mathfrak{S}'(t) \subset \mathfrak{S}(t)$ , where  $\mathfrak{S}'(t)$  excludes the winning technology  $i^*$  defined above in (3) (and quite possibly several other technologies as well). Indeed, it is a short step from here to arrive at Schumpeter's story of creative destruction if we suppose that one entrepreneur is particularly lucky or alert, perceiving  $\mathfrak{S}(t)$ , such that  $f_{i^*} \in \mathfrak{S}(t)$  while  $f_{i^*} \notin \mathfrak{S}'(t)$ . Suppose furthermore that

the entrepreneur can patent  $f_{i^*}$  and extract monopoly profits thereafter, driving every other firm from the market by cutting output price  $p$  since by (3)  $\pi_{i^*} > \pi_i$ ,  $\forall i \neq i^*$ . This monopoly will persist until some future time  $t + s$ , at which point the feasible technology set  $\mathfrak{S}(t + s)$  expands to incorporate a new optimal technology that differs from  $i^*$ . Then a new wave of creative destruction occurs.

What determines alertness to  $\mathfrak{S}(t)$ ? There is no agreement about this in entrepreneurship research. Here too, though, neoclassical economics can make some constructive suggestions. For example, Gifford (1998) considers entrepreneurs who apportion their effort between exploring new technologies and operating their existing ventures using current technology. Those entrepreneurs who devote most effort to exploration have the most extensive subset of  $\mathfrak{S}(t)$  and hence the highest probability of spotting and adopting the radical new production technology. These entrepreneurs are the most likely to be the Schumpeterian innovators, operating relatively few, possibly unprofitable, ventures using well-established technologies (Gifford, 1998). Consequently, they have the greatest incentives to explore alternative production methods outside the existing well-known possibility set  $\mathfrak{S}'(t)$ . Note that in this story entrepreneurs are still optimizing, and operating ventures using neoclassical production functions. The difference is that they are no longer passively stuck with one static production function but continually choose among them, changing technology (possibly radically) when it becomes profitable to do so. Thus Schumpeter's analysis can be seen as a natural extension of the neoclassical approach. Of course, this brief discussion has left several related questions unanswered, including whether the adoption of new and radical technologies is more or less likely at birth or at a later stage of venture development. The Industrial Organization literature on patent races and innovation is also relevant here (see below).

### 3.2. Continuous Shifts Caused by Endogenous Innovation

Prior to the 1980s, innovation was incorporated into production functions in a rather *ad hoc* manner, often by introducing a scale factor  $A(t) > 0$  to generalize (1):

$$y = A(t)f(K, L_o, L_e, M). \quad (5)$$

By making the scale factor dependent on time,  $t$ , exogenous technical change is represented, since if  $A(t)$  increases steadily over time, output  $y$  does too, holding all else equal. This was the basis of Solow's (1970) famous growth model.

An obvious drawback to the above formulation is that it fails to explain where the technical progress comes from and, in particular, the role

entrepreneurs can play in engendering and adopting innovations. Recognizing this, theorists such as Paul Romer began to endogenize innovation (see Romer, 1994, for an overview). Some endogenous growth theories allow entrepreneurs to purposively shift resources to actively pursue successful innovation strategies. For example, Aghion and Howitt (1992) studied R&D strategies when entrepreneurs can earn temporary monopoly profits from successful growth-inducing innovations. Profits are only temporary because new innovations are eventually rendered obsolete (*à la* Schumpeter) by future innovations. Aghion and Howitt showed that some degree of market power, that is, imperfect competition, is needed for Schumpeterian entrepreneurs to engage in growth-generating research. Subsequent extensions of this research program suggest that under some circumstances competition and capital can be growth enhancing too (Aghion and Howitt, 1997; Howitt and Aghion, 1998).

Another of Schumpeter's predictions was that as economies develop, large corporations exploit economies of scale in production to replace small entrepreneurial firms as the drivers of innovation. Originally framed as a verbal hypothesis by Schumpeter, recent models of purposeful R&D investment (e.g., Peretto, 1998) have been able to derive this as an outcome of optimizing behavior rather than as a general conjecture about venture development. According to a specific model by Peretto (1998), this outcome is accompanied by a prediction that new firm starts eventually peter out. Unfortunately neither of these predictions accords well with the evidence. On the former, Acs and Audretsch (1990, Ch. 2) utilized four databases on peer-reviewed "important" technological changes and innovations. Using these as measures of innovation, small firms were found to contribute around 2.4 times as many innovations per employee as large firms. On the latter, far from petering out, new firm starts and self-employment rates continue to increase in many developed countries (Parker and Robson, 2004).

One reason why new firm starts continue to occur in innovative industries may be that entrepreneurship is an effective way of bringing new ideas and production processes to market. New firms might be able to sidestep "knowledge filters" such as fixed routines or costs of changing production technology ( $\phi$  in Section 3.1 above) that inhibit established firms from exploiting innovation opportunities. Entrepreneurs can exploit these opportunities by taking advantage of "knowledge spillovers" that are obtained from knowledge production undertaken by others (Acs et al., 2004). I will cite some evidence on this issue in Section 3.4.

Researchers continue to explore how entrepreneurs innovate to improve the effectiveness of the production process rather than merely treating it as predetermined and immutable. This line of research dovetails with growing interest in the entrepreneurship literature of opportunity recognition, as well as of innovation and performance. Given the enduring interest in these topics,

this is an area that is likely to continue to attract research effort in the years to come.

### 3.3. *Idiosyncratic Entrepreneurial Ability*

Are entrepreneurs “born” or can they be “made”? This question continues to be asked of entrepreneurship scholars. If they are “born,” then one can imagine some fixed person-specific ability in entrepreneurship, denoted by a quantity  $x$ , entering the production function (1) as an additional argument. This gives rise to the “extended” production function:

$$y = f(K, L_o, x, L_e) \quad (6)$$

(suppressing the argument  $M$  for simplicity and for reasons given above). However, if entrepreneurs can be “made” rather than just being “born,” then one can think of  $x$  as being endogenous and amenable to change. While this is certainly an interesting possibility (considered briefly below), this case has been studied less than the one that takes  $x$  to be exogenous, which I concentrate on first.

It seems realistic to suppose that  $x$  is heterogeneous among the population of individuals, not all of whom will eventually become entrepreneurs. Without loss of generality, let  $x = 0$  denote the lowest ability, and let  $x = 1$  denote the greatest. Also, we have  $f(K, L_o, 1, L_e) > f(K, L_o, 0, L_e) \geq 0$ , that is, more able entrepreneurs produce more than less able entrepreneurs.

In a path-breaking paper, Lucas (1978) studied venture development among entrepreneurs who operate ventures with the production technology (6). Lucas was interested in exploring a particular aspect of venture development, namely the characteristics of those ventures that close and those that survive, as capital accumulates in the economy. Lucas showed that entrepreneurial ability is the key to understanding which firms exit and which do not, as well as what happens to the survivors. He also showed that, given any capital endowment  $K$ , only the most able individuals enter entrepreneurship. The remainder work for the entrepreneurs at a market clearing wage  $w$ . Among the entrepreneurs, those with the greatest abilities run the largest firms (where size is measured in terms of  $L_e$ ). The concavity of  $f$  ensures that there is an optimal firm size for any  $x$ , so the most talented entrepreneur does not produce everything in equilibrium. Lucas went on to show that the distribution of firm sizes reflects the (exogenous) distribution of entrepreneurial ability and that, as aggregate  $K$  increases, the entrepreneurs running the smallest ventures are the first to exit. Hence average firm size continually increases as capital accumulates. Thus Lucas predicted a perpetual decline of the small firm in the modern economy, a prediction that has

not however been generally vindicated by the evidence (see, e.g., Parker, 2004, Chapter 3).

Other extensions to the Lucas framework have also been proposed. For example, Jovanovic (1994) and Parker (2003a) recognized that ability can also affect productivity in paid employment too. In Parker's model, entrepreneurial ability not only determines productivity and occupational choice, but also entrepreneurs' access to finance under conditions of asymmetric information.

By treating  $x$  as exogenous, but determined by psychological factors, it is possible to connect the literature dealing with entrepreneurial production and occupational choice on one hand with that of entrepreneurial psychology on the other. For example, entrepreneurial ability might be embodied in a high need for achievement (McClelland, 1961), internal locus of control (Rotter, 1982), risk-taking propensity (Kihlstrom and Laffont, 1979) or tolerance of ambiguity (Timmons, 1976). It is, however, unclear whether entrepreneurs know their own ability when making decisions. Indeed, experimental evidence from Coelho (2004) suggests that individuals tend to form faulty and unrealistically over-optimistic judgments when making risky investment decisions similar to those made by entrepreneurs. The implication is that entrepreneurs make systematically inefficient production decisions. Unlike larger firms, in which decision-making is often dispersed among multiple individuals and constrained by the demands of shareholder, the owner-manager might lack internal checks on his or her flights of fancy, which might partly explain the high business failure rates of small enterprises.

Of course,  $x$  may also be amenable to change, that is, entrepreneurs might attempt to enhance their own skills to increase their productivity,  $x$ . Practical ways this might be attempted include attendance at training courses, entrepreneurship education programs and business advice. A further analysis of this issue might tie together the literature on entrepreneurship education with that dealing with entrepreneurial selection, production and performance.

To conclude, it is certainly possible to extend the neoclassical production function to incorporate idiosyncratic entrepreneurial ability. This framework can help explain who becomes an entrepreneur, as well as illuminating otherwise hidden aspects of venture development such as access to finance (Parker, 2003a). And this approach might form a natural bridge with the literature on entrepreneurial psychology and entrepreneurship education.

### 3.4. "Entrepreneurship Capital"

Finally, in this section I will briefly summarize a new hypothesis advanced by Audretsch and Keilbach (2004) relating to "entrepreneurship capital." This is the name those authors give to the collective productivity of entrepreneurs, based on a knowledge spillover story. They propose that



ventures give and receive spillovers from other ventures, for example, hi-tech firms attract a skilled labor pool that other entrepreneurs can also benefit from. In addition, formal and informal links between ventures at various stages of development can also contribute to spillovers. The aggregation of these spillovers is what Audretsch and Keilbach called entrepreneurship capital. It is denoted by  $S$  below.

There is growing evidence that knowledge spillovers exist and are economically important. For example, Acs et al. (1994) reported that innovative activity by small firms is more sensitive to university research located close to industrial research than innovation by large firms is (see also Jaffee, 1989; Audretsch and Feldman, 1996). Bearing this out, Audretsch and Lehmann (in Chapter 6 of this volume) provided evidence that knowledge- and technology-based new ventures in Germany have a greater tendency to locate close to universities, presumably in order to access knowledge spillovers. Acs and Armington (2004) reported that new firm creation rates were higher in U.S. regions that had higher proportions of adults with college degrees and a higher local density of firms. And at the more aggregate level, Acs et al. (2004) found evidence that economic growth is positively associated with the interaction between rates of entrepreneurship and R&D expenditures. Knowledge spillovers may give rise to increasing returns to scale at the aggregate level, which generates agglomeration phenomena, and which may explain persistent differences in entrepreneurship between regions (Minniti, 2005).

The relevance of entrepreneurship capital in the context of this chapter is twofold. First, it affords another extension of the neoclassical production functions given earlier of the form:

$$y = f(K, S, L_e), \quad \text{where } \partial y / \partial S > 0, \quad (7)$$

where I have suppressed the argument  $L_o$  purely for simplicity. Second, it links with the growing literature on clusters and knowledge spillovers in entrepreneurship, and the literature on endogenous growth, innovation and dynamic production functions referred to above.

Audretsch and Keilbach proposed measuring the contribution of entrepreneurship capital to entrepreneurial performance by estimating the production function (7) directly. To see how, suppose  $f$  takes the widely used Cobb-Douglas form:

$$y = A \cdot K^\alpha S^\beta (L_e)^\gamma, \quad (8)$$

where  $A$ ,  $\alpha$ ,  $\beta$  and  $\gamma$  are positive parameters. Taking natural logs of (8) and introducing a random error term  $u$  yields:

$$\ln y = \ln A + \alpha \ln K + \beta \ln S + \gamma \ln L_e + u. \quad (9)$$

Using data on a set of ventures or regions, if a proxy for  $S$  can be identified, then  $\alpha$ ,  $\beta$  and  $\gamma$  can be estimated directly. The parameter of greatest interest is  $\beta$ . If this is found to be significantly positive, then a case can be made for entrepreneurship capital to affect average entrepreneurial performance. Note, by the way, that when aggregate (rather than individual venture-based) data are utilized, it is probably admissible not to condition on  $L_o$ , on the grounds that its aggregate levels can be presumed to be fairly even across regions.

In their empirical application, Audretsch and Keilbach (2004) proposed new firm growth rates in a region as a proxy for  $S$ , and regional per capita income growth rates as a measure of performance,  $y$ . They estimated (9) using a sample of data on 327 West German regions for the year 1992. They estimated  $\beta$  as between 0.10 and 0.16, and significantly different from zero. This suggests that entrepreneurship capital is indeed a key determinant of entrepreneurial performance in Germany. Of course, the robustness of this finding remains to be verified under alternative definitions of  $S$  and using alternative sampling frames. Evidently, further research on this topic is required.

#### 4. THE ENTREPRENEUR'S OWN LABOR INPUT

A key determinant of production and performance at the level of the individual venture is the entrepreneur's own input of effort. This section reviews the current state of theoretical and empirical knowledge about:

- 4.1. The relationship between venture performance and entrepreneurs' work effort.
- 4.2 The determinants of entrepreneurs' work effort.
- 4.3 The cessation of effort in the form of retirement by the entrepreneur.

##### 4.1. Venture Performance and Entrepreneurs' Work Effort

Evidence linking entrepreneurs' work effort to their performance is sparse. This situation is often explored in the context of models of work hours. Such models shift the entrepreneur's objective from profits,  $\pi$ , to utility,  $U$ , in order to recognize the nonpecuniary (psychic) costs of supplying effort. In general, utility is taken to be an increasing function of consumption and a decreasing function of work hours which are not enjoyed for their own sake even by the self-employed (Ajayi-obe and Parker, 2005). Since consumption is constrained by current profits, one can write  $U = U(\pi, L_o)$ , where  $\delta U / \delta \pi > 0$  and  $\delta U / \delta L_o < 0$ . Differentiating with respect to  $L_o$  and setting to zero for a maximum yields the first order condition:

$$p[(\delta U / \delta \pi)(\delta f / \delta L_o)] - \delta U / \delta L_o = 0. \tag{10}$$

Define an entrepreneur's "wage,"  $v$ , as the ratio of her profits to her labor input, that is,  $v = \pi/L_o$ . Then it is possible to show (Blundell and MaCurdy, 1999) that under certain conditions, (10) can be written in the form:

$$L_o = L_o(v) = \alpha + \beta v + X \cdot \gamma + u, \quad (11)$$

where  $X$  is a vector containing several control variables, including unearned income; and  $u$  is a random disturbance term. Equation (11) is a "static" model of self-employed labor supply.

A study by the author based on U.S. PSID data finds some evidence that  $v$  is endogenous in (11), that is, that entrepreneurs' returns are directly and positively affected by their own work effort (Parker et al., 2005). The evidence for Britain is less compelling (Ajayi-obe and Parker, 2005). In both countries there is evidence that female entrepreneurs earn significantly lower wages,  $v$ , and work significantly fewer hours than their male counterparts do (Aronson, 1991). This also suggests that there is a positive link between work effort and profits. And on a different but related tack, there is also evidence that work effort can release funds from family lenders (Basu and Parker, 2001). Whether high work effort succeeds in eliciting funds from formal lenders such as commercial banks is more open to question, however. It is often difficult for banks to monitor work effort at inception and the scope for moral hazard by entrepreneurs is presumably substantial. Greater scope probably exists for venture capitalists to monitor entrepreneurs' effort, given their intensive activities of this sort (Gompers, 1995).

To date, researchers have paid more attention to the determinants of entrepreneur's work effort, rather than to the impact of effort on venture performance. Since the evidence cited above tentatively suggests that effort does affect venture performance, it is of some interest to explore those determinants, a task I turn to now.

#### 4.2. *Determinants of Entrepreneurs' Work Effort*

Previous analytical studies of entrepreneurs' work effort fall into two broad categories. One strand of research describes how entrepreneurs allocate their time between various different tasks involved in running a venture (see, e.g., McCarthy et al., 1990; Cooper et al., 1997). This includes mixing time between entrepreneurship and some other job. For example, Lévesque and MacCrimmon (1997) observed that the concavity of the production function  $f$  is necessary for work mixing to occur. The second strand of research estimates (11) in an attempt to explain the chosen quantity of work hours and the determinants of an entrepreneur's decision to retire from the workforce. The retirement decision issue is an especially topical one in view of the

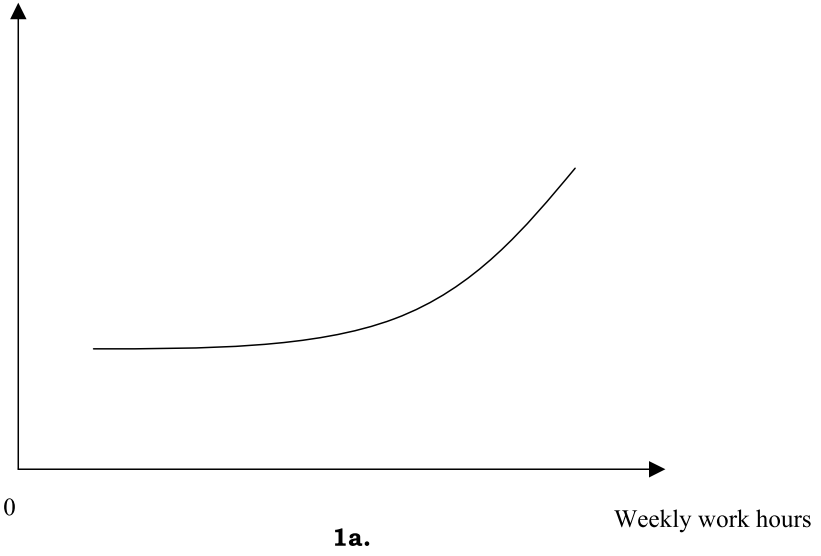
ongoing debate about promoting entrepreneurship among retirees, as a way of relieving pressure on national public pension systems. This decision is analyzed separately in Section 4.3.

What is to be gained by studying the determinants of entrepreneurs' work effort? This topic might attract policy interest for several reasons. First, a better understanding of entrepreneurs' work effort might suggest ways that current business support policies can be made to fit better with social objectives to limit work hours and foster family-friendly policies. For example, if entrepreneurs face long work hours—as the evidence suggests (Carrington et al., 1996; Hamilton, 2000; Ajayi-obe and Parker, 2005)—then policies designed to promote entrepreneurship should take account of this when, for example, targeting females who tend to have more onerous domestic commitments than males (Hundley, 2001). Second, if entrepreneurs work harder to generate higher financial returns, then there might be grounds to support income tax cuts as part of an entrepreneurship promotion policy (see Parker, 2001). However, one certainly cannot assume that entrepreneurs will always supply longer hours as  $v$  increases, a point I now explore in further detail.

It is helpful at this point to define a couple of useful terms. Economists often decompose responses of effort to wage changes into two components. These are called *substitution* and *income* effects. The substitution effect describes how individuals substitute work for leisure hours when the returns to work increase, holding their current income constant. It describes a positive relationship between hours and wages. In contrast, the income effect predicts that individuals respond to becoming better off from a wage increase by consuming more leisure and therefore working fewer hours. It implies a negative relationship between hours and wages. The total response is the sum of income and substitution effects which, being the sum of positive and negative components, is ambiguously signed in general. If the substitution effect dominates, then  $\partial L_o / \partial v > 0$  above [i.e.,  $\beta$  in (11) is positive]; and the supply of labor locus that traces out these responses will be upward sloping in effort–wage space, as illustrated in Figure 12-1a. However, if the income effect dominates at some wage  $w_1$ , say, then  $\partial L_o / \partial v < 0$  [i.e.,  $\beta$  in (11) is negative]; and the labor supply locus becomes “backward-bending” as illustrated in Figure 12-1b.

The available empirical evidence about entrepreneurs almost exclusively pertains to self-employed individuals. Wales (1973) analyzed a sample of data on American business proprietors, and estimated that most self-employed Americans were located on a “backward-bending” (rather than an upward-sloping) segment of their labor supply schedule as in Figure 12-1b. Many subsequent studies have focused on specific self-employed occupations, notably physicians, dentists and farmers. For example, using a sample of data on American physicians from the mid-1980s, Thornton (1998) found that the typical self-employed male physician was located on an upward-sloping portion

Hourly wage rate



Hourly wage rate

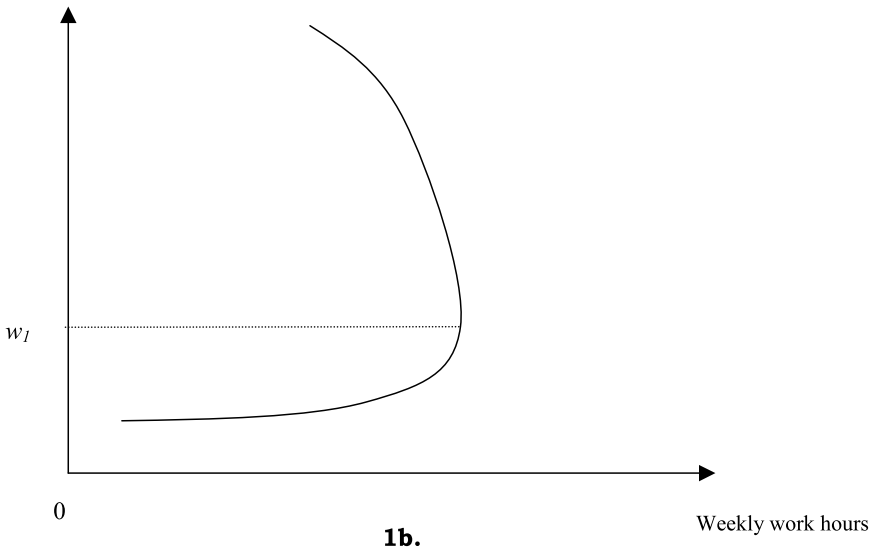


FIGURE 12-1 *Labor supply curves.*

of their labor supply curve as in Figure 12-1a. Because the gradient of the slope was steep, the labor supply responses of this group were relatively insensitive to changes in hourly wage rates and non-practice income. In contrast, Showalter and Thurston (1997) estimated statistically *significant* positive work-hours elasticities of 0.33 among self-employed physicians. In contrast, physicians in paid employment were estimated to have work hours elasticities insignificantly different from zero. In more recent work, Ajayi-obe and Parker (2005) estimated hours-wage elasticities for British self-employed workers using instrumental variable methods. For both own-account self-employed and employers, hours-wage elasticities were negative, again pointing to a dominant income effect. Estimated negative wage elasticities for self-employed American males were rendered insignificant, however, when a measure of income risk was included by Parker et al. (2005). This finding suggests that risk-averse self-employed workers may “self-insure” in response to greater risk—by working longer hours to make the deterministic part of their income larger.

In summary, while previous results from static models of self-employed work hours are mixed, there is a slight tendency for negative hours-wage elasticities to emerge for this group. One policy implication is that there may be limited scope for cutting income tax to promote work effort by entrepreneurs. In fact, self-employment entry decisions *also* appear to be fairly unresponsive to income opportunities (Parker, 2003b), a finding that could be consistent either with nonpecuniary benefits from entrepreneurship or with over-optimism about financial returns there.

A limitation of the static labor supply models considered so far is that they neglect information from other time periods that might also impact on individuals' current behavior. For example, if an entrepreneur anticipates a future recession that will decrease the demand for his good, then he might work harder now than otherwise knowing that he can take more leisure later when it is cheaper. Life cycle models incorporate information about future incomes in order to identify this kind of *intertemporal substitution*. However, future incomes and future work status are both difficult to predict for entrepreneurs—and perhaps even by entrepreneurs themselves. This might render an unmodified application of standard employee life cycle models to entrepreneurs problematic, and might in turn help explain why relatively few researchers to date have modeled entrepreneurs' labor supply behavior in a dynamic fashion. In any case, the omission of information on future entrepreneurial incomes, which are required to estimate a complete life cycle empirical model, may not be too serious if the results of Camerer et al. (1997) are more generally applicable. Camerer et al. analyzed daily variations in the work hours and wages of self-employed taxi cab drivers in New York City. They argued that cab drivers find it easier to earn money quickly on some days than on others because of exogenous factors like bad weather or train strikes. One might therefore expect cab drivers

to practice intertemporal substitution, working longer hours on high wage days and taking time off on slack days. However, their estimates of (11) yielded exactly the opposite result. Cab drivers in Camerer et al.'s sample responded to busier conditions (and hence higher hourly wages) by working *fewer* rather than more hours. This is once again indicative of a backward bending labor supply schedule. Camerer et al. concluded that it is as though cab drivers "take one day at a time," having a target income for each day, and stopping work as soon as they achieve it.

Finally, an aspect of entrepreneurial work effort that has been unduly neglected in the literature to date is that of a spouse's work effort and its contribution to production. The research that does exist tends to be from the perspective of female entrepreneurship. Thus, for example, it is known that having a self-employed husband in the household significantly and substantially increases the propensity of females to be self-employed (Caputo and Dolinsky, 1998; Bruce, 1999). In fact, 40% of self-employed American married women had self-employed spouses in 1990 (Devine, 1994). However, less is known about the extent to which spouses contribute to their partner's business, either in the form of direct labor input or indirect assistance (e.g., fielding business telephone inquiries while the partner is out on a job). For example, while trust between partners might render a spouse the most favored source of outside labor,  $L_o$ , it might not necessarily be the most productive in terms of (1). Ongoing research by the author (Parker, 2005) suggests that spouses might make similar occupational choices of entrepreneurship because knowledge spillovers tie their production functions together. More generally, further research is needed on several aspects of joint labor supply, production and performance of business owner couples.

#### *4.3. The End of Venture Development: Retirement of the Entrepreneur*

Relatively little is known about the retirement behavior of entrepreneurs apart from the fact that they tend to retire at later ages than employees do (Bruce et al., 2000; Parker and Rougier, 2004). Thus, the process of venture development can continue long into an entrepreneur's mature life, a phenomenon that might also be expected to impact on the productivity of their ventures.

What are the reasons for late retirement among entrepreneurs? In answering this question, one should distinguish between entrepreneurs who have operated businesses for many years and new entrants who switch into entrepreneurship later in their lives. According to Fuchs (1982) and Bruce et al. (2000), 2% of older Americans switch into self-employment each year compared with less than 1% in Britain (Parker and Rougier, 2004). An important point, stressed by Parker and Rougier (2004), is that the retirement behavior of "switchers" appears to be very similar to that of employees, while long-term

business owners are significantly less likely to retire than either employees or switchers. It is not yet clear what underpins the determination of some entrepreneurs at later stages of venture development to continue running their ventures into the twilight of their lives, but one factor seems to be substantial income-earning opportunities near retirement ages which delay entrepreneurs' retirement. Strikingly, none of lifetime wealth, gender, disability or poor health was found to affect retirement decisions of older British entrepreneurs in Parker and Rougier's sample, though there was some evidence that entrepreneurs were likelier to retire if their spouse did.

The scope for beneficial policy intervention in this area might be quite limited. Many of the employees that Parker and Rougier observed switching into entrepreneurship in later life had histories of low incomes and multiple job spells. Since these "switchers" were more likely to become self-employed only for short periods before retiring, the value-added of their businesses is also questionable. These results also cast doubt on the potential for generating meaningful increases in labor force participation rates via self-employment schemes for older workers.

Again, there seems to be no shortage of opportunities for further research to enrich our understanding of venture development in later life and retirement. One challenge is to model work hours jointly with actual and expected future entries into and exits from entrepreneurship through the medium of a truly dynamic framework.

## 5. PRODUCTION AND ENTREPRENEURS' PERFORMANCE

Before concluding, I will discuss briefly the relationship between production and performance. In some respects, this linkage might be deemed obvious: one might expect that more productive entrepreneurs sell more goods and services and so achieve better financial "performance." However, this does not necessarily follow. For example, some entrepreneurs specialize in not-for-profit enterprises as discussed by Helen Haugh in Chapter 14 of this volume. Such enterprises might be efficient users of production technology and highly productive in their own way, but may be observed to make negligible returns on capital from low sales whose volume is constrained by the nature of the market they are serving and their mode of organization. Another reason why a relationship between productivity and performance might not be obvious arises if one chooses other measures of venture performance. For example, survival in business can be construed as an index of "success" in its own right, in view of the high failure rates of new enterprises. Another measure of venture performance that is popular among researchers is venture growth rates. In neither case is it obvious that production and venture development map one-to-one into these outcomes.



Consider the case of survival rates first. Jovanovic's (1982) well-known model of venture entry and exit postulates that entrepreneurs learn about their productive abilities as they continue in business. The longer they trade, the better informed they are, and as time goes on only the most able entrepreneurs survive and grow their ventures. Although productivity is related to survival, this is not guaranteed. Some people can be "unlucky" in the sense that even very able entrepreneurs can receive a set of adverse signals about their ability shortly after they start up that encourages them to exit despite the fact that they would have ultimately become successful had they persisted in business. Because they exit, they never learn this.

Growth rates are also related to productivity only indirectly. The "lifestyle" choices that some entrepreneurs make to keep their ventures small and relatively under-developed (Storey, 1994) is one factor that drives a wedge between firms' productivity and growth rates. Others include disincentives to growth embodied in the tax system and the regulatory framework, and possibly also limited access to financial resources. On the other hand, Eeckhout and Jovanovic (2002) have claimed that the concave shape of  $f$  in (1) implies that small firms (which grow faster than large firms) are more productive than their larger counterparts, where productivity is measured in terms of output per worker.

Certainly empirical analyses conducted at the level of the individual firm rarely identify robust determinants of future growth that one would recognize from our earlier discussion of production technologies. A good example is the study of Westhead and Cowling (1995) in which no fewer than 67 explanatory variables were utilized in an effort to explain growth rates in small independent high-technology firms in Britain. Only a handful of statistically significant relationships were uncovered between growth and other variables, namely, firm size, the owner's education and the use of multiple sources of funding. So one cannot automatically assert a strong link between production and performance measured in terms of internal firm growth rates. Clearly, a wide range of stakeholders would be interested in the findings of any researcher who succeeded in overturning the now received wisdom that "picking winners" at the time of business inception is an impossible—and even a foolhardy—undertaking.

## 6. CONCLUSION

This chapter has discussed various aspects of "entrepreneurs as producers." The organizing theme of the chapter has been a neoclassical production function, which is a useful way of characterizing entrepreneurs' control over their environment. The production function also connects to several

other branches of the entrepreneurship literature, including entrepreneurial psychology, innovation, entrepreneurship education and female entrepreneurship. A variety of aspects of production that are specific to entrepreneurs and entrepreneurship have also been discussed, including entrepreneurial ability, the entrepreneur's choice of production functions, knowledge spillovers and entrepreneurship capital. Particular emphasis was given to the entrepreneur's own labor input, which is an essential aspect of venture development.

I would highlight two points from this overview: the extensive scope that evidently exists for further research, and the capacity of that research to speak directly to policy makers' agendas. Two future research projects are especially deserving of attention. One relates to joint household decisions to participate in entrepreneurship. As well as promising to unify several literatures relating to female entrepreneurship, labor supply and production, and venture development and performance, this topic can inform the policy debate with respect to the encouragement of female enterprises, household-level social security reforms and childcare policies. The other research agenda might investigate how entrepreneurs' work effort changes as ventures develop from start-up through to maturity. In line with the overarching theme of this volume, entrepreneurship entails not only venture creation, but also venture development and termination. As the key input into production—without which enterprises are empty shells—work effort appears to be a seriously underresearched topic in entrepreneurship, a state of affairs which, it is to be hoped, will not persist much longer.

## REFERENCES

- Acs, Z. J. and Armington, C. (2004). The Impact of Geographic Differences in Human Capital on Service Firm Formation Rates. *Journal of Urban Economics* 56: 244–278.
- Acs, Z. J. and Audretsch, D. B. (1990). *Innovation And Small Firms*. MIT Press, Cambridge, MA.
- Acs, Z. J., Audretsch, D. B., Braunerhjelm, P. and Carlsson, B. (2004). *The Missing Link: The Knowledge Filter and Entrepreneurship in Endogenous Growth*. CEPR Discussion Paper No. 4783, London.
- Aghion, P. and Howitt, P. (1992). A Model of Growth through Creative Destruction. *Econometrica* 60: 323–351.
- Aghion, P. and Howitt, P. (1997). A Schumpeterian Perspective on Growth and Competition. In *Advances in Economics and Econometrics: Theory & Applications*, Vol. 2, D. M. Kreps and K. F. Wallis (Eds), pp. 279–319. Cambridge University Press, Cambridge.
- Aigner, D., Lovell, K. and Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Models. *Journal of Econometrics* 3: 21–37.
- Ajayi-obe, O. and Parker, S. C. (2005). The Changing Nature of Work Among the Self-employed in the 1990s: Evidence from Britain. *Journal of Labor Research* 26: 501–517.
- Alvarez, R. and Crespi, G. (2003). Determinants of Technical Efficiency in Small Firms. *Small Business Economics* 20: 233–244.
- Aronson, R. L. (1991). *Self Employment: A Labor Market Perspective*. ILR Press, Ithaca, NY.

- Audretsch, D. B. and Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86: 630–640.
- Audretsch, D. B. and Keilbach, M. (2004). Does Entrepreneurship Capital Matter? *Entrepreneurship Theory & Practice* 28: 419–429.
- Basu, A. and Parker, S. C. (2001). Family Finance and New Business Start-ups. *Oxford Bulletin of Economics & Statistics* 63: 333–358.
- Baumol, W. J. (1968). Entrepreneurship in Economic Theory. *American Economic Review Papers & Proceedings* 58: 64–71.
- Blundell, R. and MaCurdy, T. (1999). Labor Supply: A Review of Alternative Approaches. In *Handbook of Labor Economics*, Volume 3A, O. Ashenfelter and D. Card (Eds), pp. 1559–1695. Elsevier, Amsterdam.
- Bruce, D. (1999). Do Husbands Matter? Married Women Entering Self-employment. *Small Business Economics* 13: 317–329.
- Bruce, D., Holtz-Eakin, D. and Quinn, J. (2000). *Self-employment and Labor Market Transitions at Older Ages*. Working Paper CRR WP No. 2000-13, Boston College, Boston, MA.
- Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997). Labor Supply of New York City Cabdrivers: One Day at a Time. *Quarterly Journal of Economics* 112: 407–441.
- Caputo, R. K. and Dolinsky, A. (1998). Women's Choice to Pursue Self-employment: The Role of Financial and Human Capital of Household Members. *Journal of Small Business Management* 36: 8–17.
- Carrington, W. J., McCue, K. and Pierce, B. (1996). The Role of Employer/Employee Interactions in Labor Market Cycles: Evidence From the Self-employed. *Journal of Labor Economics* 14: 571–602.
- Carroll, R., Holtz-Eakin, D., Rider, M. and Rosen, H. S. (2000). Income Taxes and Entrepreneurs' Use of Labor. *Journal of Labor Economics* 18: 324–351.
- Coelho, M. P. (2004). *Unrealistic Optimism, Entrepreneurship and Adverse Selection*. PhD Dissertation, London School of Economics, London.
- Cooper, A. C., Ramachandran, M. and Schoorman, D. (1997). Time Allocation Patterns of Craftsmen and Administrative Entrepreneurs: Implications for Financial Performance. *Entrepreneurship Theory & Practice* 22: 123–136.
- Cowling, M. (2003). *The Contribution of the Self-employed to Employment in the EU*. Report to the Small Business Service, URN 03/539, <http://www.sbs.gov.uk/content/research/ContributionSelf-Emp.pdf>.
- Cowling, M., Mitchell, P. and Taylor, M. (2004). Job Creators. *Manchester School* 72: 601–617.
- Devine, T. J. (1994). Characteristics of Self-employed Women in the United States. *Monthly Labor Review* 117: 20–34.
- Eeckhout, J. and Jovanovic, B. (2002). Knowledge Spillovers and Inequality. *American Economic Review* 92: 1290–1307.
- Fuchs, V. R. (1982). Self-employment and Labor-Force Participation of Older Males. *Journal of Human Resources* 17: 339–357.
- Gifford, S. (1998). *The Allocation of Limited Entrepreneurial Attention*. Kluwer, Boston.
- Gompers, P. (1995). Optimal Investment, Monitoring, and the Staging of Venture Capital. *Journal of Finance* 50: 1461–1489.
- Hamermesh, D. S. (1993). *Labor Demand*. Princeton University Press, Princeton.
- Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-employment. *Journal of Political Economy* 108: 604–631.
- Howitt, P. and Aghion, P. (1998). Capital Accumulation and Innovation as Complementary Factors in Long-Run Growth. *Journal of Economic Growth* 3: 111–130.
- Hundley, G. (2001). Why Women Earn Less than Men in Self-employment. *Journal of Labor Research* 22: 817–829.

- Jacoby, H. G. (1993). Shadow Wages and Peasant Family Labor Supply—An Econometric Application to the Peruvian Sierra. *Review of Economic Studies* 60: 903–921.
- Jaffee, A. B. (1989). Real Effects of Academic Research. *American Economic Review* 79: 957–970.
- Jefferson, P. N. (1997). Unemployment and Financial Constraints Faced by Small Firms. *Economic Inquiry* 35: 108–119.
- Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica* 50: 649–670.
- Jovanovic, B. (1994). Firm Formation with Heterogeneous Management and Labor Skills. *Small Business Economics* 6: 185–191.
- Kihlstrom, R. E. and Laffont, J. J. (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy* 87: 719–749.
- Kirchhoff, B. A. (1991). Entrepreneurship's Contribution to Economics. *Entrepreneurship Theory & Practice* 16: 93–112.
- Kirzner, I. (1985). *Discovery and the Capitalist Process*. University of Chicago Press, Chicago.
- Kuhn, P. J. and Schuetze, H. J. (2001). Self-employment Dynamics and Self-employment Trends: A Study of Canadian Men and Women, 1982–1998. *Canadian Journal of Economics* 34: 760–784.
- Lévesque, M. and MacCrimmon, K. R. (1997). On the Interaction of Time and Money Invested in New Ventures. *Entrepreneurship Theory & Practice* 22: 89–110.
- Lucas, R. E. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics* 9: 508–523.
- McCarthy, A. M., Krueger, D. A. and Schoenecker, T. S. (1990). Changes in the Time Allocation Patterns of Entrepreneurs. *Entrepreneurship Theory & Practice* 15: 7–18.
- McClelland, D. C. (1961). *The Achieving Society*. Van Nostrand, Princeton, NJ.
- Meyer, B. (1990). *Why Are There so Few Black Entrepreneurs?* NBER Working Paper, No. 3537, Cambridge, MA.
- Minniti, M. (2005). Entrepreneurship and Network Externalities. *Journal of Economic Behavior & Organization* 57: 1–27.
- Moralee, L. (1998). Self-employment in the 1990s. *Labor Market Trends* 106: 121–130.
- Parker, S. C. (2001). Risk, Self-employment and Differential Taxation. *Manchester School* 69: 1–15.
- Parker, S. C. (2003a). Asymmetric Information, Occupational Choice and Government Policy. *Economic Journal* 113: 861–882.
- Parker, S. C. (2003b). Does Tax Evasion Affect Occupational Choice? *Oxford Bulletin of Economics & Statistics* 65: 379–394.
- Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Parker, S. C. (2005). *Entrepreneurship among Married Couples in the United States: A Simultaneous Probit Approach*. Discussion Paper No. 1712, IZA Institute for the Study of Labor, Bonn.
- Parker, S. C. and Rougier, J. (2004). *Explaining the Retirement Behavior of the Self-employed*. Working Paper No. 04/08, Department of Economics & Finance, Durham University.
- Parker, S. C. and Robson, M. T. (2004). Explaining International Variations in Self-employment: Evidence from a Panel of OECD Countries. *Southern Economic Journal* 71: 287–301.
- Parker, S. C., Belghitar, Y. and Barmby, T. (2005). Income Uncertainty and the Labor Supply of Self-employed Workers. *Economic Journal* 115: C190–C207.
- Peretto, P. F. (1998). Technological Change, Market Rivalry, and the Evolution of the Capitalist Engine of Growth. *Journal of Economic Growth* 3: 53–80.
- Praag, C. M. van and Cramer, J. S. (2001). The Roots of Entrepreneurship and Labor Demand: Individual Ability and Low Risk Aversion. *Economica* 68: 45–62.

- Romer, P. M. (1994). The Origins of Endogenous Growth. *Journal of Economic Perspectives* 8: 3–22.
- Rotter, J. B. (1982). *The Development and Applications of Social Learning Theory: Selected Papers*. Praeger, New York.
- Schumpeter, J. A. (1934). *The Theory of Economic Development*. Harvard University Press, Cambridge, MA.
- Showalter, M. H. and Thurston, N. K. (1997). Taxes and Labor Supply for High-Income Physicians. *Journal of Public Economics* 66: 73–97.
- Solow, R. M. (1970). *Growth Theory: An Exposition*. Oxford University Press, Oxford.
- Storey, D. J. (1994). *Understanding the Small Business Sector*. Routledge, London.
- Thornton, J. (1998). The Labor Supply Behavior of Self-employed Solo Practice Physicians. *Applied Economics* 30: 85–94.
- Timmons, J. A. (1976). Characteristics and Role Demands of Entrepreneurship. *American Journal of Small Business* 3: 5–17.
- Varian, H. (1992). *Microeconomic Analysis*, 3<sup>rd</sup> edition. Norton, New York.
- Wales, T. J. (1973). Estimation of a Labor Supply Curve for Self-employed Business Proprietors. *International Economic Review* 14: 69–80.
- Westhead, P. and Cowling, M. (1995). Employment Change in Independent Owner-Managed High-Technology Firms in Great Britain. *Small Business Economics* 7: 111–140.

PER DAVIDSSON<sup>a</sup> , LEONA ACHTENHAGEN<sup>b</sup> and LUCIA NALDI<sup>c</sup>

<sup>a</sup>*Brisbane Graduate School of Business, Queensland University of Technology (and Jönköping International Business School), Gardens Point, Brisbane, 4001 Queensland, Australia*

*E-mail: per.davidsson@qut.edu.au*

<sup>b</sup>*Jönköping International Business School, Sweden*

*E-mail: leona.achtenhagen@jibs.hj.se*

<sup>c</sup>*Jönköping International Business School, Sweden*

*E-mail: lucia.naldi@jibs.hj.se*

## 13. What Do We Know About Small Firm Growth?

### 1. INTRODUCTION

Almost twenty years ago it struck the first author of this chapter that small firm growth (and nongrowth) of small firms could be a suitable topic for his doctoral dissertation project. Seconds later he concluded that this topic probably was exhausted already. However, this turned out not to be the case. Instead, it was surprisingly hard to find any studies at all that had the growth of small and/or young and/or owner-managed firms as their main focus. From the few studies that did, and adding a somewhat larger number of studies that at least marginally touched upon the issue, a rather patchy and confusing image emerged. It was clear that we did not have much systematic, generalizable knowledge about the phenomenon of small firm growth (Davidsson, 1989a).

Today it is no longer true that studies of small firm growth are short in supply. On the contrary, as demonstrated by reviews by Ardishvili, Cardozo, Harmon and Vadakath (1998), Delmar (1997), Storey (1994) and Wiklund (1998), dozens and dozens of empirical research studies on this topic can be compiled. Does that mean we know all we need to know about the growth of small firms? Not necessarily, as all of the reviewers just mentioned complained that a coherent picture is not easy to distill from the material. This is likely due to differences in theoretical and epistemological perspectives and interpretations; operationalizations; empirical contexts; modeling and analysis approaches, as well as the inherent complexity of the phenomenon itself. Thus, not only a superficial but also a rather deep reading of the

extant literature easily leaves the reader confused and wondering. Admitting that, we will focus in this chapter on the fact that significant progress has been made and that we do actually know quite a bit now about the phenomenon of small firm growth; about its antecedents and effects, and about how it can or should be studied. It is not possible within the confines of a book chapter—and is possibly outside the capacity of these authors—to give a complete account of that knowledge. What we will attempt is a brief summary of points of convergence within some key themes. We will focus on the growth of the company or organization rather than the expansion of specific business activities although—as we shall see—these different notions of growth often coincide in the case of small firms. Admittedly, while literature from the U.S. and to some extent other parts of the world will also be used, a (Western) European perspective will dominate our treatment of the topic.

After first discussing the nature of the phenomenon and its relation to entrepreneurship, we will move on to how growth can best be assessed. A major section, composed of several sub-sections, is devoted to findings on factors that contribute to or hinder firm growth. Following that we offer a section on *how* small firms grow, if at all. In particular, we discuss organic growth vs. acquisitions, as well as growth through internationalization. The next topic we treat is “growth stages and transitions.” This overlaps with several of the issues dealt with in other sections but as it represents a relatively separate stream in the literature we keep it as such. Before concluding, we also treat the effects of growth in terms of profitability and job creation. We choose to focus on these two aspects as they arguably represent the most important outcomes on the firm and societal levels, respectively.

## 2. WHAT IS GROWTH?

### 2.1. *Growth as Process and Change in Amount*

In discussing what firm growth is we find it wise to consult the only true classic in this area, Edith Penrose. In her seminal book she characterizes the phenomenon of growth as follows: “The term ‘growth’ is used in ordinary discourse with two different connotations. It sometimes denotes merely increase in amount; for example, when one speaks of ‘growth’ in output, export, and sales. At other times, however, it is used in its primary meaning implying an increase in size or improvement in quality as a result of a *process* of development, akin to natural biological processes in which an interacting series of internal changes leads to increases in size accompanied by changes in the characteristics of the growing object” (Penrose, 1959: 1). This distinction is important for the remainder of this chapter. Most research has undoubtedly been

directed at explaining differences in the *amount* of growth and neglected other aspects of the process of growth. The primary exception is the literature on stages-of-development (or organizational life cycles) where consequences of the process of growth are a key theme.

## *2.2. The Heterogeneity of Growth*

Our use of the concept in this chapter will reflect this previous use in the literature, that is, the size-change perspective will dominate. Even if restricted to one of these conceptualizations, however, growth remains a multi-faceted phenomenon. For example, Delmar, Davidsson and Gartner (2003) discussed heterogeneity according to what specific measure the firm grows and also as regards the appropriateness of these different measures relative to specific theories. They further treated heterogeneity in the regularity or irregularity of growth over time, and in the type of growth (organic or acquisition based). Empirically, they show that when the top 10% “high growth firms” in a large sample of firms was singled out according to six different growth indicators, over 40% qualified according to at least one criterion. However, only 16.6% made the hurdle for three or more criteria and a tiny 2.5% were classified as “high growth firms” regardless of what criterion was used. Underlying this are very low correlations between some of the growth indicators, as these researchers also reported. By means of cluster analysis they distill seven different types of “high growth firms” which show markedly different growth patterns (and background characteristics). They conclude that firm growth is a multidimensional rather than one-dimensional phenomenon and that different forms of growth may have different determinants and effects. Consequently, they may also need different theoretical explanations (cf. Davidsson and Wiklund, 2000).

In addition to what Delmar et al. (2000) discussed, growth can also take different forms in terms of vertical integration, related or unrelated diversification, or be achieved through modes like licensing, alliances or joint ventures (Killing, 1978; Levie, 1997; Roberts and Berry, 1985). We will not be able to consistently consider all this heterogeneity throughout this chapter. The prototypical growth firm we have in mind unless otherwise stated is one that experiences relatively stable growth in sales over considerable time, and where this growth in sales is at least to some extent accompanied with accumulation of employees and assets, so that organizational and managerial complexity increases with growth.

## *2.3. How Does Growth Relate to Entrepreneurship?*

Firm growth is undoubtedly a topic of relevance within many branches of economics and management studies. As this manuscript appears in an



entrepreneurship context the relationship between growth and entrepreneurship is of particular relevance. Some hold that “growth is the very essence of entrepreneurship” (Sexton, 1997: 97) or make differences in sales growth the criterion for distinguishing between entrepreneurial and nonentrepreneurial firms (Birch, 1987; McDougall, Covin, Robinson and Herron, 1994). Davidsson (1989a) argued that, to the extent the owner-manager has a choice, going for growth is more entrepreneurial than not doing so when both alternatives are feasible, just as starting a firm is considered more entrepreneurial than not doing so. Davidsson, Delmar and Wiklund (2002) later delved more deeply into the growth-entrepreneurship relationship and arrived at the following: If entrepreneurship is understood as the creation or emergence of new organizations (Gartner, 1988; Gartner and Carter, 2003), growth is not formally part of the definition of the phenomenon (cf. Meyer, Neck and Meeks, 2002). However, as most start-ups remain one-person businesses, or at least very small for their entire existence (Aldrich, 1999; Reynolds and White, 1997), it makes sense to include what others would call “early growth” because otherwise entrepreneurship research cannot fill the gap between nonexistence and established organizations as we know them from organization studies. If entrepreneurship is instead defined as *creation of new economic activity* or some close alternative to that (Low and MacMillan, 1988; Shane and Venkataraman, 2000; Stevenson and Jarillo, 1990), firm growth is an aspect of entrepreneurship if it is achieved through the introduction of new products or services. If it consists solely of demand-driven volume expansion for existing products or is achieved through the acquisitions of business activities that were already up and running within another organization, growth is not an aspect of entrepreneurship (cf. Davidsson, 2003, 2004). We will be able to uphold these distinctions here only to the extent that the design of the reviewed studies so allows.

The link between entrepreneurship and growth is also relevant when considering the theory of the firm in which both entrepreneurship and growth play important roles. In Cassons’s view (2000: 116), “the modern theory of the firm addressed four main issues: [...] the boundary of the firm; the internal organization of the firm; the formation, growth and diversification of the firm; and the role of the entrepreneur.” He maintains that the role of the entrepreneur is the most fundamental. Similarly, in Penrose’s (1959) theory of the growth of the firm, entrepreneurship is the *conditio sine qua non* of continuous growth. As Penrose (1959, p. 8) writes: “For a firm, enterprising management is the one identifiable condition without which continued growth is precluded—this is one necessary (though not sufficient) condition for continued growth.” Specifically, it is the continuous exploitation of new productive opportunities which drives the growth of the Penrosian firm.

### 3. HOW SHOULD GROWTH BE ASSESSED?

#### *3.1. Assessing Change in Amount*

Even if growth is viewed merely as change in amount, it is inescapable that this change occurs over time. This means that firm growth should be researched longitudinally at least in the sense that assessment of the predictors precedes assessment of the outcome, that is, the change in size. Although the use of longitudinal designs is increasing, a large number of previous growth studies were in fact cross-sectional. This means that researchers have been involved in “prediction of the past” or have made strong assumptions about causal order and/or nonchangeability of the predictors over time. Cross-sectional studies assessing growth from an earlier point in time up to the time of the investigation are also subject to selection (success) and hindsight (retrospection) biases. Hence, further empirical contributions to this literature ought to employ a longitudinal design.

From the change-in-amount perspective, growth can be measured with a range of different indicators, the most frequently suggested being sales, employment, assets, physical output, market share and profits (Ardishvili et al., 1998; Delmar, 1997; Weinzimmer, Nystrom and Freeman, 1998; Wiklund, 1998). In within-industry studies, even more specialized measures are conceivable, such as the number of seats for restaurants or theaters, and the number of vehicles for taxi or car rental companies (Bolton, 1971). Among available alternatives the researcher has the choice to: a) create a multiple indicator index; b) use alternative measures separately; and c) find the one best indicator. If growth is conceived of as a latent construct with common causes but alternative manifestations the multiple-indicator index makes sense (Davidsson, 1991). The underlying theory here is that the same explanatory factors facilitate or hinder growth across firms, but that this growth for some firms manifests itself as, for example, radically increased sales turnover without much change in assets or employment, whereas for other types of firm the result is moderate and balanced growth across, for example, assets, employment and sales. The sum of standardized versions of all three indicators would then be a better representation of the theoretical growth concept. If only one indicator were used, results would be weak and possibly distorted.

Alternatively, the underlying theory predicts that certain antecedent would be related to, for example, growth in sales and market share while other predictors are believed to influence growth in employment and profits, respectively. If so, the sensible course of action is to include and analyze different growth indicators separately (Delmar, 1997). The theoretical and empirical evidence is leaning in favor of this other notion. For example, Chandler, McKelvie and Davidsson (2005) successfully used transaction cost

theory to explain when growth in sales and employment do and do not move closely together.

Another defensible alternative would be to confine the study to the one growth indicator that is best matched with the theory or perspective in question. If only one indicator is used and the study has a cross-industry design there is growing consensus that sales growth should be the preferred choice (Ardishvili et al., 1998; Hoy, McDougall and Dsouza, 1992; Weinzimmer et al., 1998; Wiklund, 1998). It is the most general of the alternatives as all commercial firms need to have sales to survive. According to Barkham, Gudgin, Hart and Hanvey (1996), it is also the indicator small firm owner managers use themselves. (It could be remarked, though, that both researchers and managers themselves use sales as a proxy for firm value due to the unavailability of data on that measure until the firm is actually traded.) In addition, it may be argued that sales often precede the other indicators; it is the increase in sales that necessitates increases in assets and employees and results in rising profits or market share (Flamholtz, 1986). These favorable aspects of sales as indicator is reflected in that with 30.9% of the studies it is the most used in research reviewed by Delmar (1997). Almost as popular is employment growth, which was the choice in 29.1% of the reviewed studies. While the most relevant for some purposes such as policy makers' interest in fostering employment growth through entrepreneurship (Davidsson and Wiklund, 2000), this indicator is probably often applied for reasons of data availability. Very few managers see growth in employees as a goal in itself (Gray, 1990; Wiklund, 1998; Robson and Bennett, 2000) and because some growing firms outsource heavily employment growth is not always highly correlated with sales growth (Delmar et al., 2003).

The other indicators are less generally applicable and therefore not applied as frequently. The "market" in market share calculations may be ambiguous; differences in market share may be irrelevant for small firms, and comparing shares for firms operating in different markets may be indefensible. The value of assets varies with the capital intensity of industries and is difficult to assess where the key asset is knowledge. Physical output can hardly be compared across industries. While profits are universally relevant they reflect many other aspects of a firm apart from its size. Besides, it is perfectly possible for a large and/or growing (in sales or employment) firm to be unprofitable (Davidsson, Steffens and Fitzsimmons, 2005).

While sales may be the most universally applicable growth indicator, it is not always the best one. As Penrose (1959: 199) stated almost half a century ago, "there is no way of measuring an amount of expansion, or even the size of a firm, that is not open to serious conceptual objections." For example, high-tech companies with rather long development times, such as biotech companies, are not able to display any growth in sales or revenues for long periods of time. Yet, during this period they might still grow in terms of assets—including

knowledge assets such as patents—and employment. In other cases, the revenue figure may be inflated by one-off divestment of business units rather than only capturing sales of products and services. When data cover several countries and/or time periods, differences in inflation rates are a complicating factor. Moreover, it has recently been argued that employment rather than sales growth is the indicator that has the most consistent positive correlation with other growth measures (Wiklund, 2005). Rather than using sales because others have proposed it, researchers are well advised to think seriously about what growth indicator(s) best matches their theory, their research questions, and the type of firms included in their own sample.

The distinction between organic growth and growth through acquisitions has been widely ignored in previous research (Delmar et al., 2003). When the key interest of the study is on the societal level, this is a crucial distinction as acquisition-based growth in itself does not bring any net addition to the economy. Also, in studies on the firm level, this distinction deserves more scrutiny as the drivers and effects of the two forms of growth are likely to have differential managerial implications (Levie, 1997; Penrose, 1959). Therefore, when possible it seems a wise decision for researchers to choose a data collection procedure that allows them to partial out organic from total growth.

Apart from choice of indicator the specific formula used to calculate growth may affect results. This is additional reason to include and analyze different indicators separately so as to detect and make sense of such differences (Delmar, 1997; Weinzimmer et al., 1998). In particular, it has been observed that effects of firm size on growth vary depending on whether an absolute or a relative measure is used. In short, relative (percentage) measures tend to “favor” small firm growth while the reverse is true for absolute growth measures. It may be argued that sophisticated researchers have no problem understanding this complication and that the inclusion of size as a control variable solves the problem. While it does in a technical sense, a range of other independent variables may be size-dependent in nonobvious ways so that also their estimated effect on growth is sensitive to whether an absolute or a relative growth measure is used. Therefore, the size-sensitivity of specific formulae deserves deeper consideration than the mere inclusion of size as a control variable.

Further, the use of only first-year and end-year data for growth calculations has been criticized because it models growth as one giant leap (Davidsson and Wiklund, 2000) and makes the calculation overly sensitive to stochastic variation (Weinzimmer et al., 1998). On this ground, the latter suggest that the slope of the regression line over multi-period data be used as the measure of firm growth. To some extent such a practice also narrows the gap between the size change and process perspectives on growth.

### 3.2. *Assessing Growth Processes*

However, strong research on growth as process calls for a fundamentally different approach. The arms-length, quantitative study of determinants of growth does not put much flesh on the bone to understanding the issue from a process point of view. This can create a major challenge as a number of the determinants fostering or hindering growth are not stable over time. Attitudes and motivation of founders/CEOs could, for example, change dramatically due to events in their business or private lives. A classic example in the literature is Stanworth and Curran's (1973) "Frank Williams" case. Wiklund (1998: 87) discussed the difficulty as follows: "...we really do not know how much variables change over the studied time period, and whether or not this is a major problem. Growth, as such, is a change process and it could be that explanatory variables change quite substantially during this process. Until we do know, it must remain an unwise oversimplification to assume that nothing else but size changes." While existing studies manage to give an answer to the question of how different determinants affect growth, they largely fail to explain the underlying processes of why these determinants might affect growth.

When growth is conceived of as a process, there is little doubt that having several indicators of growth is preferable and that these need to be assessed at several different points in time. Especially if the study is of a close-up nature, a very rich image can be captured, including, for example, direct assessment of organizational complexity along several dimensions as the growth process unfolds. This is not to say processes cannot be studied quantitatively. However, it requires considerable resources and staying power on the part of the research team to study a substantial number of development processes in an intense manner (Raffa, Zollo and Caponi, 1996). While retrospective reconstructions of growth processes do not lack value, they are subject to potential biases due to hindsight and rationalization after the fact on the part of informants. To some extent this can be remedied with use of multiple informants and documents produced at the time, but whether qualitative or quantitative in nature, a more ideal study would follow the growth processes as they evolve.

## 4. WHAT FACTORS FACILITATE OR HINDER GROWTH?

### 4.1. *Internal vs. External Determinants*

First, it is important to realize that growth is not the norm. Most firms start small, live small and die small (see Chapter 16 of this volume). They never embark on a significant growth trajectory (Aldrich, 1999; Reynolds and White, 1997; Storey, 1994). One major reason for this is that the majority of start-ups are imitative businesses in mature industries serving local markets (Aldrich,

1999; Reynolds, Bygrave and Autio, 2003; Samuelsson, 2001, 2004). As such, they do not have much growth potential.

Potential or not, it is also clear that most business founders have modest growth aspirations for their firms. This has been demonstrated in several different studies across countries (Cliff, 1998; Delmar and Davidsson, 1999; Dennis and Solomon, 2001; Human and Matthews, 2004). But does the manager's willingness to grow really matter, or do external forces largely determine the firm's growth, as suggested by the population ecology perspective (Hannan and Freeman, 1977)? Environments vary along dimensions such as dynamism, heterogeneity, hostility and munificence (Dess and Beard, 1984), and these external factors rather than the managers' motivations and strategic actions may largely determine how much the firm grows. For example, it has been clearly demonstrated that rapidly growing firms are more often found in industries and regions that are more dynamic (Carroll and Hannan, 2000; Davidsson and Delmar, 1997, 2001; Jovanovich, 1982). While in highly innovative industries the failure rate for new entrants is also higher, Audretsch (1995) demonstrated that for those who survive the first few years both survival and growth is higher in subsequent years for firms in more innovative industries. Growth firms in industries that are stagnant overall are often found in dynamic growth niches within these industries (Storey, 1997; Wiklund, 1998). This seems to correspond to Penrose's (1959, p. 222ff) discussion of the opportunities for small firms to enter and grow in a market which she calls the *interstices* in an economy. These are productive opportunities which small firms see and believe they can take advantage of that are left open by the large firms.

The growth effects of other dimensions of environments are less well established. While confirming the positive effect of dynamism (in his case *increase* in dynamism), Wiklund (1998) found a weak negative effect of environmental hostility, and no effect of heterogeneity. It is likely that these other environmental conditions are associated with contradictory effects so that the overall effect can be zero or tilt over in either direction depending on the specific context. For example, resource munificence may facilitate the building of capacity to grow but also attract more new entrants that compete for the market potential for growth. It has been argued that in heterogeneous markets, entrepreneurial opportunities are more likely to arise as developments in one market creates demand for a firm's products in related areas (Zahra, 1991). However, heterogeneity may also indicate that the market is fragmented into small niches across which individual firms would find it difficult to expand.

Thus, the evidence suggests that firm growth is to a certain extent externally determined. On the other hand, studies that include explanations on different levels tend not to highlight environmental characteristics as being the most influential (Davidsson, 1991; Wiklund, 1999). Although the odd study may have failed to establish such a relationship (e.g., Jenkins and Johnson,

1997) there is also compelling evidence that the owner-manager's growth motivation, communicated vision and goals have direct effects on the firm's growth (Baum and Locke, 2004; Baum, Locke and Kirkpatrick, 1998; Delmar and Wiklund, 2003; Kolvereid and Bullvåg, 1996; Mok and van den Tillaart, 1990; Wiklund, 2001; Wiklund and Shepherd, 2003) as has the firm's strategic orientation (Bamford, Dean and McDougall, 1997; O'Gorman, 1997; Wiklund and Shepherd, 2005).

Taken together, the sensible conclusion is that growth is to a considerable extent a matter of willingness and skill, but that fundamental facilitators and obstacles in the environment cannot be disregarded. The extent to which the firm governs its own destiny is also likely to vary across firms and situations. For example, the image that emerged from Davidsson and Delmar's (1997, 2001) research is that firms in the subgroup they define as high growth find ways to reach their growth goals relatively regardless of environmental conditions, while the majority of "other firms" seems to swing up and down with the development of the economy at large. Over a deep recession and recovery, the "other firms" in their study first markedly decreased and then increased employment. Since they are defined on that basis it is no surprise that the curve for "high growth firm" was located much higher on the growth axis and never hit negative numbers. The compelling feature, however, is that there was no downturn at all for this category of firm. A closer look reveals that this was achieved by increasing the amount of acquisition-based growth in hard times; like other firms, the high-growth firms are largely unable to expand organically under such conditions.

It should also be noted that it is sometimes difficult to determine what factors are truly "external" and "internal," respectively. For example, in Chandler and Hanks' (1994) conceptualization, qualities of the opportunity are regarded as aspects of the environment. Davidsson (1989a, 1991) likewise associates the opportunity concept with environmental factors. In more recent works, "opportunity" is often used interchangeably with "business idea" and interpreted as an internal issue (cf. Davidsson, 2003, 2004).

#### 4.2. *The Influence of Selected Growth Determinants*

Compiling mostly U.K. studies from the late 1980s and early 1990s, but without combining them in an integrated model, Storey (1994) organized the evidence in the categories *the entrepreneur*, *the firm*, and *strategy*. Support for influence is found in all three categories. Among the variables associated with the individual, a majority of studies found that for *motivation*, *education*, *management experience*, *number of founders* and *functional skills*, the influence on growth is positive, although the last factor had only been investigated in two studies. Variables associated with the individual entrepreneur are also at the heart of Jovanovic's (1982) model which assumes that individuals have different

innate abilities but imperfect information about them when starting a business. A particularly interesting feature of his model is that entrepreneurs learn about their true abilities as the business survives and grows. *Unemployment* as start-up reason was mostly negatively associated with growth, whereas for prior self-employment, social marginality (ethnicity), training, age, prior sector experience and gender the evidence was mixed or most studies suggested they had no effect on growth (see also Chapter 16).

Some of these generalizations deserve elaboration. The positive effect of team size (number of founders) has been rationalized as different team members making up for each other's competence deficits, that is, a diversity argument (Cooper, Gimeno-Gascon and Woo, 1994). Yet, Ruef, Aldrich and Carter (2003) found team composition to be driven by similarity, not diversity. Ensley, Pearson and Amason (2002) found that top manager team cohesion in new ventures is actually positively related to new venture growth. One reason for this might be that past joint work experience among the founding team members increases their speed in decision making, as proposed by Eisenhardt and Schoonhoven (1990).

The lack of a gender effect is also important to comment. This is one of the more certain generalizations, as the variable was included in most of the studies Storey (1994) reviewed. Other research suggests that women-owned businesses do not seem to under perform with regard to profitability, employment or orders (DuRietz and Henrekson, 2000). When studies suggest that female-owned businesses grow less (e.g., Cooper, Gimeno-Gascon and Woo, 1994; Fischer, Reuber and Dyke, 1993), it is likely to be either an industry effect rather than a true gender effect, or a result of lower average growth aspirations on the part of female business owners, indicating neither less effective use of resources nor lesser ability to reach their goals (Cliff, 1998; DuRietz and Henrekson, 2000; Watson, 2002).

As regards the firm, the evidence suggests that *firm age and size, sectoral affiliation, legal form* and *location* are all systematically related to growth. As regards size, all studies found a significant effect but the sign varies, probably as a consequence of the specific growth measure employed (cf. above). Note that some factors here attributed to the firm coincide with what has above been discussed as environment, and that Storey's (1997) compilation largely confirms what was stated there. Especially the discussion of age and size as determinants of firm growth has a long tradition, following the formulation of Gibrat's law in 1931. Gibrat's law states that the rate of growth of a firm is independent of its size at the beginning of the period, and that the probability of a given growth rate during a specific time interval is the same for any firm within the same industry. However, empirical studies typically do not find support for the independence of firm growth from size and age (Becchetti and Trovato, 2002: 291).



As regards strategy variables, the evidence is much less conclusive than for the firm variables. For variables that were included in five or more studies, a relatively consistent positive effect was found for *technological sophistication*, *market positioning* and *new product introduction*. In individual studies, several other strategy variables were also shown to be influential but collectively the evidence was weak, mixed, or the factor had been included in too few studies for any conclusions to be drawn.

Regarding innovation, the argument that large enterprises are the driving engine of innovative activities has been questioned by several researchers (Acs and Audretsch, 2003). There is evidence that at an aggregate level SMEs spend less on R&D than large firms but produce almost twice as many innovations on a per-employee basis (Acs and Preston, 1997). Evidence provided below on small growth firms' tendency to expand organically is in line with the notion that their growth is often innovation-based.

Expanding the strategy issue beyond Storey's (1994) review we find that in a comprehensive, longitudinal study, and combining strategy and human capital arguments, Raffa et al. (1996) found that firms based initially on technical entrepreneurial know-how expand their market abilities by 1) collaborating with large firms, 2) collaborating with professionals and consultants, 3) using external (technical and market) competencies and 4) acquiring new market competencies through diversification of the entrepreneurial group's activities or new market-oriented employees. In contrast, firms initially based on strong entrepreneurial market knowledge faced more difficulties in supplementing their know-how with technical skills. We have also noted above that Entrepreneurial Orientation (i.e., innovativeness, pro-activeness and risk-taking) positively affects growth. Some caution is recommendable, though, as it has been shown that the different sub-dimensions of EO may have differential effects on firm performance (Lumpkin and Dess, 2001). Wiklund and Shepherd (2005) were also able to demonstrate that the effect of EO—in this case on an index combining growth and financial performance—is moderated by environmental dynamism and capital availability. This is direct evidence that strategy needs to be adapted to the environment and a likely reason why few findings on strategy are generalizable across many studies. This may also explain why some studies arrive at counterintuitive results on strategy. For example, both Bamford, Dean and McDougall (1997) and McDougall et al. (1994) found that broad strategies were more successful with respect to small firm growth, thus questioning the otherwise common niche argument.

The existence of contingencies and interaction effects also points at where research on firm growth stands today. Rather than assuming linear, additive effects, research increasingly focuses on fit and combined effects. Representing different disciplines, Chandler and Hanks (1994) and Audretsch (1995) were both forerunners in this trend. There are several reasons for this

development. Generally increased methodological sophistication of this field of research is one, probably fueled by disappointment over relatively weak results in many earlier studies. Increased theory-drivenness is another, especially as there has also been a shift from theories that regard firms as essentially similar microunits (Hannan and Freeman, 1977; Porter, 1980) to those that emphasize their uniqueness (Barney, 1991, 1997; Wernerfelt, 1984, 1995). The use and usefulness of analysis of moderators is not limited to strategy variables. While Storey (1994) found mostly positive effects of education and management experience, others have emphasized that these effects are surprisingly weak (e.g., Davidsson, 1989a). The reason for the latter is easy to understand in light of moderation results reported by Wiklund and Shepherd (2003). They found that education and experience have much stronger relation to growth if growth aspirations are also high. That is, ability gained through experience and education does not deterministically force business founders to expand their firms. If they aspire to do so, however, education and experience seem instrumental in reaching that goal.

#### 4.3. *Integrated Models of Growth Determinants*

Evidently, many different internal and external factors could under some circumstances affect firm growth, and consequently a very long list of specific growth determinants has been suggested in the literature. This poses a challenge for studies aiming at approaching full explanation of the phenomenon of small firm growth, rather than testing effects predicted by a particular theory. On the one hand, it has to include a broad range of explanatory variables; on the other hand, some abstracted sense-making is needed, that is, the grouping of the many specific variables under a smaller number of over-arching themes. While also other individual studies cover a range of factors on different levels (e.g., Eisenhardt and Schoonhoven, 1990; Sandberg and Hofer, 1987), Davidsson (1991) and Wiklund (1998) represent two out of few attempts to formally integrate a broad range of growth determinants in a causal model and to test it empirically.

In Davidsson's model, all low-level specifics are regarded as aspects of three exhaustive factors: *ability*, *need* and *opportunity*. He further distinguished between objective and perceived versions of these variables, but as his study was cross-sectional, only the objective factors could be related to actual growth in the empirical analysis. His results show that all three factors affect growth but also that the variables indicating variance in the *need* for growth were the most influential. They also had the most stable effects across industries. The same pattern emerged when objective and perceived ability, need and opportunity were related to future growth aspirations.

Wiklund combined three theoretical perspectives in his model: the *resource-based view*, the *motivation perspective* and *strategic adaptation*. In his model, strategy—operationalized as Entrepreneurial Orientation; *EO* (Lumpkin and Dess, 1996)—is hypothesized to be directly related to growth, whereas resources, motivations and characteristics of the environment are assumed to indirectly affect growth via strategic adaptation. His results confirm that all included categories of variables influence growth. However, in empirical estimation, aspects of motivation and the environment were ascribed direct effects alongside their effects via strategy. Subsequent analyses have shown that the *EO*–performance link increases in strength over time, at least over periods of moderate length (Wiklund, 1999). Taking this into consideration, his results support the notion that strategy has the strongest and most direct influence on growth. This is an important addition to Davidsson's (1991) conclusions as explicit consideration of strategy was lacking in his study.

While Davidsson's and Wiklund's models captured many factors and distinguished between indirect and direct effects, they did not include interactive (or moderated) effects which were shown above to be important. Doing both at the same time may be beyond the capacity of any researcher, or even the statistical software used. An alternative strategy is then to confine the study to one level of analysis (or one disciplinary perspective) and to limit other influences by drawing a sample from a relatively homogeneous empirical context. An excellent and recent example is Baum and Locke's (2004) psychological study of determinants of firm growth. Confining their study to a population of North American architectural woodwork firms and including a small number of firm- and environment level control variables, these researchers found strong direct effects of goals, communicated vision and self-efficacy on growth over a six-year period. In line with their theory, they also found mostly indirect effects of passion, tenacity and new resource skills. In a less carefully operationalized study, and using a more heterogeneous sample, these relationships may well have been undetected.

#### 4.4. Growth Barriers

Barriers to growth are to a considerable extent the same as the mirror image of the drivers of growth that have been discussed above (Barber, Metcalfe and Porteous, 1989). However, certain factors are more commonly discussed from the perspective of their negative influence. Examples include various institutional factors. Noting that indisputable evidence for the effects of institutional arrangements is almost impossible to establish, Davidsson and Henrekson (2002) hold that the consistency of the theoretical arguments and empirical data makes a strong case for the notion that in the case of post WWII Sweden, certain institutions have systematically discriminated against the growth of independent

businesses. The specific institutions they investigated included, for example, regulation of certain sectors of the economy, taxation, wage-setting institutions and labor market legislation. Carlsson (2002) employed a broader perspective on institutions in his comparison of technology clusters in Sweden and Ohio. The factors he investigated include the science base, mechanisms for technology transfer, density of networks and what he calls “entrepreneurial climate.” Again, the conclusion is that Swedish institutions have hampered firm growth.

Carlsson’s (2002) study also included capital availability and the author pointed out this as one of the institutional factors particularly likely to explain differential growth patterns for firms in Sweden and Ohio. This may well be the case with respect to the type of firms his study included. There are also other studies that have pointed at provision of external debt and equity capital as a very important factor in promoting small firm growth (e.g., Becchetti and Trovato, 2002; Riding and Haynes, 1998). However, it is naïve to conceive of the economy as populated by small firms that are all full of willingness and potential to grow if only the financial means were available. We will not attempt full coverage of this complex and thorny issue here. Penetrating this topic quite thoroughly, Storey (1994) arrived at the conclusion that there is no general market failure that motivates a major role for government in improving the financing of small firms. As regards private external capital, the issue is loaded with motivational concerns, agency problems, procedural justice issues and possible detrimental effects of overfunding (Cressy and Olofsson, 1996; Davidsson et al., 2005; Sapienza, Korsgaard and Forbes, 2003; Wiklund, Davidsson and Delmar, 2003). For these reasons also those firms that face profitable growth opportunities may refrain from growth or go for growth only if they can do so based on retained earnings or financial bootstrapping (Winborg and Landström, 2001). The issue is far more complex than just being a matter of providing enough external capital for these firms that have growth potential but lack the resources to realize it.

## 5. MODES OF SMALL FIRM GROWTH

### 5.1. *Organic Growth vs. Acquisitions?*

In response to this question it is important again to remember that most firms do not grow. Among those that do, Davidsson and Delmar (1998), who investigated the entire population of Swedish firms that had 20 or more employees in 1996 and backtracked their development for up to ten years, demonstrated that small (and young) firms had a much stronger tendency to grow organically than large firms had. As these specific results have not previously been published in a widely accessible outlet, it makes sense to repeat them in full here, as shown in Table 13-1. The table includes data only for firms

TABLE 13-1 *Total and organic growth for high growth firms of different sizes*

| End year size class | No. of cases ( <i>n</i> ) | Cumulative total employment growth | Cumulative organic employment growth | Organic as percent of total |
|---------------------|---------------------------|------------------------------------|--------------------------------------|-----------------------------|
| 20–49               | 342                       | 8124                               | 7963                                 | 98.0                        |
| 50–249              | 532                       | 44 320                             | 34 208                               | 77.2                        |
| 250–499             | 127                       | 22 340                             | 12 497                               | 55.9                        |
| 500–2499            | 127                       | 57 752                             | 15 682                               | 27.2                        |
| 2500+               | 25                        | 52 728                             | –10 310                              | (–19.6)                     |
| Total               | 1153                      | 185 264                            | 60 040                               | 32.4                        |

Source: Davidsson and Delmar (1998).

that have first been classified as being among the top 10% growers in absolute, average annual growth in employment. For members of this growth elite, the table breaks down their expansion according to what proportion is organic and achieved through acquisition, respectively.

The difference in growth mode by size class is quite dramatic. In the smallest size class almost all growth is organic. This share then drops monotonically and sharply across size classes. In the largest size class firms that are classified as “high growth firms” based on growth in total employment, actually shrunk quite dramatically in organic terms. The same clear pattern is repeated if the breakdown is made by beginning-year size class or by age. Among high-growth firms that are five years or younger the organic share is about 90%, whereas among those that are older than ten years only 16% of the growth is organic.

Not many other studies have explored this relationship. McCann (1991, p. 191) argued that dominance for internal venturing among young and relatively inexperienced firms is not surprising as such firms hardly have the resources to grow aggressively via acquisitions (cf. Penrose, 1959; Wiklund and Davidsson, 1999). Empirically, Kraemer and Venkataraman (1997) focused on firms that possessed inventions at start-up and found that these were more likely to venture internally than through acquisitions or strategic alliances. In a more broadly based study (albeit restricted to manufacturing firms) of young, growing firms in France, Ireland and Scotland, Levie (1997) obtained results similar to Davidsson and Delmar’s (1998) although size and age differences are not quite as dramatic in his study. This is probably partly due to the fact that his study excluded all firms that have fewer than 50 employees. Levie’s study also explored diversification and integration strategies. The results revealed that the great majority of firms grow in volume within a single industry or engage in related diversification. Very few firms engage in vertical integration or unrelated diversification. While volume growth and some

related diversification dominated the picture, Levie's data suggest a select minority of high growth firms utilize a broader range of growth modes. This resonates with Killing (1978) and Roberts and Berry (1985) who suggested that licensing, alliances and joint ventures are important for high growth firms. Accordingly, Barringer and Greening (1998) found that about half of the firms in their sample of high growth firms had engaged in strategic alliances. It is possible, however, that their focus on geographical expansion led to a high estimate.

### 5.2. Growth through Internationalization

The issue of alliances leads us to networks and, as we shall see, to growth through internationalization. The role of networks has long been a prominent topic in entrepreneurship research, both in the discussion of entrepreneurs' personal networks (e.g., Birley, 1985) and firm networks (e.g., Butler and Hansen, 1991). A number of studies explicitly link networks to firm growth (Donckels and Lambrecht, 1995; Hansen, 1995; Jarillo, 1989). A network perspective has also been applied in different studies of firm internationalization. For example, Chetty and Campbell-Hunt (2003) investigated the relationships between rapid international growth and business networks. In their study, business networks were the only vehicle for internationalization out of a small domestic market in a sudden internationalization process, when large increases in capabilities were involved and when it involved specialization. The link between networking and internationalization is built on establishing and maintaining the required relationships with business partners, customers, suppliers and governments (Welch and Welch, 1996).

Networks or not, there is a growing body of literature which aims at understanding firm growth through *internationalization*. Due to today's low-cost, rapid world-wide communication and transportation, the domain in which firms operate and expand is becoming truly international (McDougall and Oviatt, 1997). This appears especially true for small countries (Julien and Ramangalahy, 2003). Thus, the globalization of markets and the consequent need for crossing national borders does not only concern large and established firm (Bloodgood, Sapienza and Almeida, 1996). "Internationalization is no longer regarded as an alternative but rather as an essential prerequisite for growth, also for small businesses' "(Hurmerinta-Peltomäki, 1994, p. 24).

The international expansion of small and medium sized firms is regarded as an entrepreneurial act since it entails the opening up of product markets (Thorelli, 1989; Ibeh, 2003). The same is true for any geographic expansion. Interestingly, however, geographic expansion is almost exclusively discussed in the context of internationalization. Notable exceptions are the

studies by Barringer and Greening (1998) and Greening, Barringer and Macy (1996). The former argued (1998, p. 490) that opening a new geographic site is similar to a start-up process in that a firm must select a location, hire and train staff, establish organizational legitimacy, motivate and supervise employees, and develop a structure to accommodate future growth. This resonates well with Davidsson's (2003, 2004) argument that geographical expansion is a form of entrepreneurship although he, like Thorelli (1989) and Ibeh (2003), rather emphasizes the similarity with the start-up situation from the perspective of the market effects.

Researchers interested in international entrepreneurship have focused mainly on what has been labeled "international new ventures" (INVs) (McDougall, Shane and Oviatt, 1994), "high growth new ventures" (Bloodgood et al., 1996) or "born globals" (Madsen and Servais, 1997). These are "new entrepreneurial ventures with high aspiration and potential for growth" (Bloodgood et al., 1996) and "business organization that from inception, seek to derive significant competitive advantage from the use of resources and the sale of output in multiple countries" (McDougall et al., 1994, p. 153). Autio, Sapienza and Almeida (2000) found that the earlier in their development firms venture into international competition and the greater their knowledge intensity, the more rapidly these firms expanded internationally.

Much of this research has been presented as an alternative to the view provided by studies that explain internationalization as a gradual process, occurring through a sequence of stages (Johanson and Vahlne, 1977; Johanson and Wiedersheim-Paul, 1975). The latter approach has been criticized as being too deterministic and stressing only the early stages of internationalization (Melin, 1992). Yet, unlike most research on firm growth, the literature acknowledges that the process is not always unidirectional. Rather, studies also show how firms reduce their international activities or withdraw from international operations (e.g., Benito and Welch, 1997); how they withdraw from foreign direct investment and return to exporting (Chetty, 1999), or—less dramatically—drop single products or product lines (Calof and Beamish, 1995).

While highlighting an important form of expansion, the research on international growth of new and small firms has so far not yielded many strong generalizations. In part this may be due to the even greater complexity involved in such research when the development of firms serving home markets of vastly different size is compared. In part, it may also be due to this stream of research not yet having achieved the same level of theoretical and methodological sophistication as the best exemplars of the "determinants of growth" research.

## 6. GROWTH STAGES AND TRANSITIONS

### 6.1. Overview of the Literature

Apart from attempts at finding growth facilitators and obstacles there exists a whole body of literature, which is more concerned with the processes of growth. This type of research is often presented in the form of life cycle or stages models that encompass the entire life span of an organization (e.g., Adizes, 1989; Churchill and Lewis, 1983; Greiner, 1972; Hanks, Watson, Jansen and Chandler, 1994; Flamholtz, 1986; Galbraith, 1982; Quinn and Cameron, 1983; Churchill and Lewis, 1983; Scott and Bruce, 1987; Kazanjian, 1988, and many more). These models attempt to provide a more dynamic view on the development of organizations and their growth (cf. Aldrich, 1999: 196–201). Usually, *life-cycle models* abstractly represent a cycle of emergence, growth, maturity and decline. Whetten's (1987) work on organizational growth and decline is an example of this type of research, as well as Adizes's (1989) model which distinguishes between the growing and the aging sides of the life-cycle curve. Interestingly, a number of these life-cycle models—while discussing growth at the organizational level—implicitly maintain that these changes over the history of an organization would be the manifestation of a similar population-level phenomenon (O'Rand and Krecker, 1990). Other models explicitly test life-cycle processes in population-level studies of organizational mortality and/or founding (Hannan and Freeman, 1987; Carroll and Delacroix, 1982). Here, the occurrence of Stinchcombe's (1965) concept of the 'liability of newness' is often hypothesized. Freeman and Carroll (1983), for instance, found empirical support for the assumption that new organizations are more likely to cease trading than are old organizations. At the same time, their results show that besides the liability of newness, the liability of smallness also concurs in explaining firms' rate of dissolution. However, in the literature, decline and death of organizations have received somewhat less attention than birth and early development (Neumair, 1998). Indeed, the vast majority of models considers mainly the firm's development process up to the maturity stage and frequently focuses on the generic problems organizations encounter during growth. These have been referred as developmental or *stage models*. Firms are assumed to grow in distinct stages, each stage concluded by a set of typical problems and organizational responses.

The number of stages and sub-stages identified by the scholars varies significantly (O'Farrell and Hitchens, 1983). All models start with an initial stage which is typically characterized by a simple organizational structure, direct supervision, and particular importance is attributed to the founder or entrepreneur: for example, Greiner's (1972) "creativity stage;" Churchill and Lewis' (1983) "existence stage;" Quinn and Cameron's (1983) "entre-



preneurial stage;” Kazanjian and Drazin’s (1989) “conception and development stage;” and Adizes’ (1989) “infant stage.” In the following stage, the firm achieves its initial product market success (Miller and Friesen, 1984). Here, a first division of managerial tasks occurs, but control is still achieved through personal supervision (O’Farrell and Hitchens, 1983). This stage corresponds to Greiner’s (1972) “direction stage,” to Churchill and Lewis’s “survival” and “success” stages, to Kazanjian and Drazin’s (1989) “commercialization” stage; to Adizes’s (1989) “go-go stage,” and Garnsey’s (1998) “resource generation.” The subsequent stages are characterized by an increased bureaucratization of the organizational structure and by the separation between management and control: for example, Churchill and Lewis’ “resource maturity” and Quinn and Cameron’s (1983)’s “formalization and control stage.”

In a related fashion, there is a literature on growth transitions and typical managerial growth problems which does not necessarily discuss a set number of stages that firms are assumed to go through (Arbaugh and Camp, 2000; Fombrun and Wally, 1989; Hambrick and Crozier, 1985; Hofer and Charan, 1984). There are also examples of contributions that point out some positive outcomes of the growth process itself. For example, Rollag (2001) argues that rapid growth helps to socialize the employees into a venture more rapidly.

## 6.2. Critique and Further Developments

Stages or life-cycle models are, on the one hand, intuitively appealing as they directly address the issue of new venture growth and accurately point at the gradual nature of firm evolution. However, on the other hand, they only conform to a uniform path of growth in a deterministic way (e.g., Fombrun and Wally, 1989). They build on assumptions that organizations pass through all the stages of the life cycle and that there would be an optimal configuration for each stage (cf. Wiklund, 1998). In reality, young ventures, for example, might simply experiment with new organizing principles within the same stage, and these would not be accounted for. In addition, stages models are cyclical in the sense that they do not tend toward equilibrium but rather return to a starting point (cf. Stubbart and Smalley, 1999). Life-cycle models in particular show the process as primarily dependent on the time factor. In other words, organizations follow the same time consistent pattern as they grow and decline (Hofer and Charan, 1994). A further point of criticism is that the models mainly focus on the evolving of formal structures, though it is well known that informal structures and processes (such as the informal networking of the entrepreneurial team) are of great importance (Birley and Stockley, 2000). Thus, the models oversimplify the nature of the role of the entrepreneur or entrepreneurial team.

Their motivation, decisions and actions have a great impact on the growth process but are hardly considered in these models. The models also imply that managerial action should be narrowly prescribed if growth is to occur (Tang, Jones and Forrester, 1997).

In addition, many of the models share the problem of lacking systematic empirical evidence (Gibb and Davies, 1990). A growth model that fares better in that regard is Hanks et al. (1994). Explicitly setting out to tighten the life-cycle concept, these researchers cluster analyzed a sample of 126 high-technology organizations in Utah in order to establish whether distinct development stages could be discerned empirically and, if so, which they were. These authors found four clusters that correspond to development stages of increasing complexity and first increasing and then decreasing dynamism, respectively. The different clusters also differ as regards firm age and a range of internal characteristics. What makes their results even more realistic, however, is that they found another two clusters that did not fall naturally into a stages model. These were firms that either never had entered into a path of dynamic development or those that had more or less permanently left such a path. Hence, the Hanks et al. (1994) categorization responds to the criticism of previous models being overly deterministic and lacking systematic empirical backing (cf. also Churchill and Lewis, 1984).

The Hanks et al. (1994) study is subject to limitations such as being based on one particular industry and inferring transitions through stages from age differences in a cross-sectional analysis. Admitting this, theirs is definitely one of the most rigorous attempts toward a research-based stages model. Ironically, the popularity of stage models seems to have declined dramatically since its publication, much like Woo, Cooper and Dunkelberg's (1991) critical examination of "types of entrepreneurs" seems to have made that research stream peter out. This is unfortunate as research-based knowledge on growth processes and transitions would have high practical relevance alongside research findings on growth facilitators and obstacles. Process knowledge can make entrepreneurs aware of possible crises and solutions, and researchers should be able to present better alternatives to the portrayals of inevitable growth problems and universally applicable snake oil cures that one finds in the nonresearch-based management literature. One of few recent efforts in this research stream is Garnsey's (1998) attempt to extend Penrose's work to early growth (Penrose is mainly concerned with established firms). Garnsey explicitly discusses growth reversal or stability as common growth paths. Unfortunately, even though she acknowledges that it would be important to understand the micro processes of growth (1998, p. 551), Garnsey also stays at an abstracted level, thus making her findings less directly relevant for managers.

## 7. THE EFFECTS OF GROWTH

### *7.1. Desirable and Undesirable Effects of Growth*

Both in academic and nonacademic literature, firm growth is frequently equated with success (cf. Baum, Locke and Smith, 2001; referring to Covin and Slevin, 1997, and Low and MacMillan, 1988). However, as pointed out in the growth stages and transitions literature reviewed above, growth can lead to a number of undesirable consequences or “growing pains” (Flamholtz and Randle, 1990). Small firm owner-managers are generally aware that growth can have both desirable and undesirable effects, and hence growth is something of a dilemma for them. In research directly addressing small firm owner-managers’ expectations as to the negative and positive consequences of growth, it has been found that expectations of economic gain is not a dominant growth motivator, that almost all respondents expect both negative and positive outcomes and that negative expectations are overall somewhat more frequent or pronounced than positive ones (Davidsson, 1989b; Wiklund et al., 2003). The strongest dominance for negative expectations concerned the issue of vulnerability; a majority believed that increased size would make their firms less able to survive a severe crisis. This is likely a misconception as the bulk of evidence suggests a positive relationship between size or growth on the one hand and survival on the other (Aldrich and Auster, 1986; Davidsson, Lindmark and Olofsson, 1998; Stinchcombe, 1965; Storey, 1994).

Wiklund et al. (2003) further showed that consistently across three separate studies and various subsample breakdowns, the strongest negative effect on overall growth willingness stems from expectations that growth would have adverse effects on employee well-being, which they interpret as fear of losing the informal, family-like character of the small organization. As regards this concern, the research literature lends some support to the owner-manager’s fears: small organizations have certain advantages that risk being lost if the organization grows larger (Arrow, 1983; Barker and Gump, 1964; Mosakowski, 2002). As mentioned above, many owner-managers also resent the idea of achieving growth based on substantial influx of external capital (Sapienza et al., 2003). Clearly, then, small firm owner-managers expect growth to bring both positive and negative outcomes, and they are not all wrong in doing so.

The following section will discuss two outcomes in more detail, namely, profitability and an increase in number of employees. Arguably, the former is one of the most important potential effects of growth for the (owner-) managers of firms, while the latter represents a key interest among policy makers.

### 7.2. Is Growth Profitable?

Regarding the relationship between growth and profitability, Davidsson's (1989b) research showed that 40% of the small firm owner-managers in his sample did not believe growth would improve their personal income stream, thus effectively removing one important reason to pursue growth. While fairly strong theoretical arguments can be put forward both for growth enhancing profits and for profits enhancing growth, the fact is that the research evidence on the association between growth and profitability is surprisingly weak and mixed. One might have thought that this issue was settled once and for all when, on the basis of a meta-analysis of 320 studies published in 1921–1987, Capon, Farley and Hoenig (1990, p. 1148) concluded that "Growth, analyzed in 88 studies, is consistently related to higher financial performance. Growth in assets and sales individually show positive relationships to performance at both industry and firm/business levels of analysis." However, a close examination of their analysis (Table 5; p. 1154) discloses that a significant positive effect of growth on financial performance is found in across-industry studies only. In within-industry studies, the effect is minuscule in magnitude and statistically not significant. This is actually evidence *against* the hypothesis that firms that grow more than their close competitors become more profitable as a result.

It is surprisingly difficult to find more recent studies that explicitly examine the growth-profit relationship. Chandler and Jansen (1992), Mendelson (2000) and Wiklund (1998) all found a positive association in passing; their main research questions concerned other relationships. A few recent studies have addressed the growth-profitability as their main research question. Cox, Camp and Ensley (2002) surveyed 672 members of the Entrepreneur of the Year Institute and found a positive relationship between sales growth rate and profitability growth. Cowling (2004) investigated U.K. firms across industries and concluded from a series of regression analyses that profit and growth tended to move together. However, Roper (1999), who studied a large sample of Irish firms, found turnover growth and return on assets to be very weakly related ( $r$  below 0.10 and not statistically significant). Likewise, Sexton, Pricer and Nenide (2000), who analyzed over 75,000 firms in the Kauffman Longitudinal Financial Statement Database, found a very weak overall correlation between sales growth and profitability. Markman and Gartner (2002) used longitudinal data on Inc. 500 firms and found that change in sales and change in employment both had a weak *negative* correlation with *change* in profit.

Hence, the empirical evidence on the relationship between growth and performance is inconclusive. In addition, to the extent a relationship exists it has not been determined whether this is primarily because growth leads to profits or, conversely, because profitability drives growth. This triggered Davidsson et al. (2005) to recently examine precisely that question. Their results showed that firms originating in the high profit/low growth category were in each analysis

about two to three times more likely to end up in the desirable high growth/high profit category as were firms originating in the high growth/low profit category. The latter category was instead strongly overrepresented among firms regressing to a low profit/low growth position. This is strong reason to caution against a universal and uncritical growth ideology and for small firm owner managers—whenever possible—to secure a sound level of profitability before they go for growth. While perhaps appropriate under some circumstances, as a general rule the idea of growing in order to become profitable seems a much more questionable prospect.

### 7.3. Firm Growth and Job Creation

From a societal point of view the creation of new jobs—resulting in increased tax revenue and reduced welfare costs—is often the vantage point for an interest in firm growth. There is no doubt that the majority of gross new jobs in the economy are the result of growth of already existing firms, rather than entry of new firms. In the case of Sweden, the proportion has been estimated as roughly one third for entry and two thirds for expansion (Davidsson, Lindmark and Olofsson, 1998). This should come as no surprise as there are many more established firms in an economy than there are new entrants. The more important question concerns where net additions of jobs come from. As noted by Davis, Haltiwanger and Schuh (1996a, 1996b) this is a trickier issue, as *post hoc* a given net addition can be attributed to many different subcategories of the economy. Studies in the U.S. and U.K. have claimed that a small minority of rapidly growing firms—so-called “flyers” or “gazelles”—are the real creators of net new jobs in the economy (Birch and Medoff, 1994; Birch, Haggerty and Parsons, 1995; Storey, 1994). Studies in Sweden have not been able to find a minority of gazelles that sum up to impressive absolute numbers of new jobs (Davidsson and Delmar, 1997, 2003). On the contrary, the entry and early, modest growth of a large number of “mice” seems to be the major source of net new jobs in Sweden (Davidsson et al., 1996, 1998).

The differences in results may in part be due to real country differences. For example, the small home market in a country like Sweden may lead to smaller numbers of firms that grow really big. Alternatively, the firms that do so move abroad or at least their expansion occurs in other countries and may be concealed from the figures available to the researcher. However, to a certain extent the notion that a small number of high growth firms are responsible for a very large share of employment gains can be the result of a method artifact. As demonstrated by Davidsson (2004, pp. 160–163), if one follows a cohort of firms over time and there is any outcome variance at all—even completely stochastic variance—it will always be the case that a small proportion of firms eventually accounts for a large proportion of the jobs created *by that cohort*.

The greater the outcome variance and the longer the analysis period, the more marked will this effect be. However, this does not prove that the elite of high-growth firms create a large proportion of all new jobs *in the economy*. In order to establish the latter, the job creation of *all gazelles* in the economy has to be compared with total job creation in the economy.

There are also other reasons not to equate employment growth on the firm level with job creation in the economy at large. As noted above, especially for large and old firms, growth usually reflects acquisition, that is, transfer of already existing activities of jobs from one organization to another. Even those firms that grow organically may do so at the expense of other firms whose employment consequently shrinks. Yet other firms contribute to the growth of the economy by reducing the need for manpower for a given output. Clearly, head counting on the firm level is a very narrow sighted analysis for societal purposes. When the interest truly is in the size of employment in the economy and its changes it seems advisable to start at a more aggregate level and then try to tease out—on region, industry and firm levels—how the aggregate effects emerge from entry, exit, expansion, contraction and transfer of economic activities across borders.

## 8. CONCLUSIONS

Our review has demonstrated that small firm growth is a complex phenomenon. The concept “growth” denotes both a change in amount and the process by which that change is attained. Further, the growth can be achieved in different ways and with varying degrees of regularity, and it manifests itself along several different dimensions such as sales, employment and accumulation of assets. This complexity has naturally led researchers to adopt different approaches to studying growth and to use different measures to assess it. Further, although our review shows that it can fruitfully be regarded as a growth issue, the research on small firms’ internationalization has largely developed as a separate stream. Similarly, other relatively separate literatures have evolved, which effectively focus on different modes of growth although mostly without regarding the studies first and foremost as growth studies. This goes for topics like acquisitions and mergers, diversification and integration—research streams which have largely ignored the particularities of small firms and which in turn have been largely ignored among researchers focusing on small firm growth.

Despite this complexity and fragmentation, a considerable body of generalizable knowledge about small firm growth now exists, which is what we have tried to highlight in this chapter. One could easily emphasize the problems instead: weak conceptualization of the phenomenon of organizational growth; lack of integration of the different findings into a more comprehensive theory

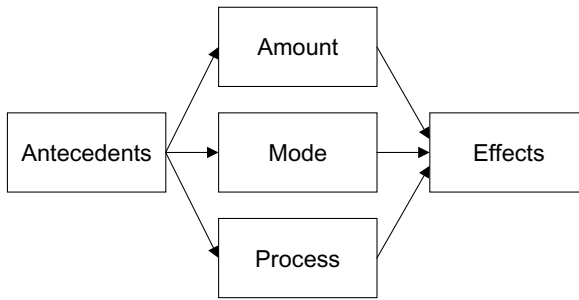


FIGURE 13-1 *Alternative foci for studies of small firm growth.*

of growth; lack of high quality in-depth studies; focus on unidirectionality of growth; rather weak links between empirical findings and theory-building, etc. However, the luxury of seeing such deficiencies can only be enjoyed because many researchers put considerable effort into researching firm growth, thus little by little uncovering the true complexity of the phenomenon. What previous research—and taking stock of it—has achieved more than anything else is to clarify what aspects of this complex phenomenon have been relatively well researched and which remain virtually virgin ground. The remaining validity of some of the criticism of previous research only means that there are interesting research opportunities for followers to do better.

So what are these research opportunities? We choose to organize our discussion of future research needs around Figure 13-1.

Let us turn first to the question how *Antecedents* relate to the *Amount* of growth. This is the firm growth subtopic which consciously or not has attracted the most interest in previous empirical research. We would hold that there is little need for studies that try to identify factors that facilitate, predict or hinder growth. There are enough such factors identified in the literature already; the likelihood that any important ones would have been neglected is slim. Neither do we think it very meaningful to further explore the relative importance of different factors for the growth of “small firms in general.” The population of small firms is too heterogeneous for this to be a very meaningful exercise, and the effects probably too country-, cohort- and period-specific for such results to have much theoretical value. At least we would think investing in new and comprehensive empirical studies has limited value until one has first taken more systematic stock of the knowledge that is already available. This could take the form of conducting formal meta-analyses of the research that is already available, including the assessment of moderators (e.g., specific growth measure used) and applying various theoretical tools to attain a deeper understanding of the meaning of the results such an exercise unveils.

When this groundwork has been properly done, it is conceivable that comprehensive empirical studies of the growth of “small firms in general”

would have great value. A well-designed study of that kind should probably combine the strategies of applying a high level of abstraction and paying attention to the interplay between different influences as discussed in our above review. However, we find it likely that a more fruitful way forward is to conduct theory-driven studies of growth within more homogeneous samples of firms. Baum and Locke's (2004) psychological study is an exemplar in this regard. Using homogeneous samples is a way of controlling for the otherwise often confounding influence of variables one does not have a theoretical interest in (Kish, 1987). Moreover, the use of homogeneous samples allows one to use operationalizations that are maximally relevant for the particular type of firm or industry. The issue of broader generalization, we would hold, is better dealt with through replication across several different samples, each of which is internally relatively homogeneous, than by trying to include all different types of small firms in the same study (cf. Davidsson, 2004).

Let us now turn to studies of how the *Amount* of growth relates to various *Effects*. A general observation here is that in studies of firm growth, some type of positive "ultimate" effects of growth are often implicitly or explicitly assumed without being tested. As implied by our above review of research on the effects of growth, we believe it is time for researchers to do better than just assuming that firm growth is an end in itself. When growth is the dependent variable used, researchers should explain on what grounds, how and for whom they believe firm level growth to be important. We also need more studies of how growth relates to important management level goals such as profitability and firm value. The effects of growth in terms of various management challenges that have to be dealt with are an issue we will return to shortly.

In policy-motivated research, it is frequently assumed that head counting on the micro level translates to corresponding employment effects on aggregate levels of analysis. Phenomena such as acquisition-based growth and one firm's growing at the expense of others suggest that such an assumption is overly simplistic. For the purpose of policy as well as for testing industry- or region-level theory, the relationship between firm growth and economic development is better studied by letting the prevalence of high growth organizations compete with other measures of economic dynamism for explaining aggregate level economic development (Davidsson et al., 1996, 1998).

Different *modes* of growth are a clearly underresearched area in the small business literature. It is so underresearched, in fact, that studies which merely map out the phenomenon would have considerable value even if they say nothing about antecedents and effects. The question of modes would also benefit from increased integration of the knowledge that has been gained in already existing literatures that relate to growth. Internationalization is one such example which we have here made an attempt to integrate with the



growth literature. Other such areas of theorizing and empirical research, which clearly can inform our knowledge of growth whether framed in that way or not, are those dealing with diversification and integration, even if they so far have rarely dealt with small firm issues specifically. Research and theorizing concerning acquisitions have not typically focused on the problems of small firms, either; nor have they always portrayed the phenomenon as a growth issue to be compared with other modes of growth. However, the few empirical attempts that have been made to investigate issues of modes of small firm growth (other than internationalization) have yielded some very interesting results that certainly deserve a follow-up. Studies are needed that can confirm or call in question, for example, Davidsson and Delmar's (1998) result that there is a very strong relationship between (small) firm size and the tendency to grow organically, or Levie's (1997) observation that a distinct minority of rapidly growing firms display an array of different modes of achieving growth. Still more importantly, a theoretical understanding of what such relationships mean needs to be developed. For example, do small firms grow organically because they are more innovative or because they lack the resources to choose the (safer and sounder?) acquisition route to increased size?

We noted early in this chapter that Penrose (1959) pointed out that "growth" does not only mean "change in amount." It sometimes also denotes the *process* by which this change comes into being. This is a sorely underresearched area and therefore another one where mere mapping of the phenomenon has value, although relationships with antecedents and effects are, of course, of the greatest interest. While the "stages-of-development" or "life-cycle" literature can be rather elaborate on process issues, and while considerable commonalities exist across many such accounts, the empirical evidence is not impressive. The quantitative material, when existing at all, is typically cross-sectional and retrospective. To the extent concurrent process data underlie the theorizing, it is often gathered rather unsystematically from the small, nonrandom sample of firms for which the theorist happens to have consulted.

What is needed here are case-based studies where the cases have been sampled on sound, theory-based criteria, as well as quantitative work on samples that are known to likely represent some relevant population of firms. These studies would need to avoid retrospection bias and the "prediction of the past" of cross-sectional research by studying concurrent growth processes with a longitudinal design. Preferably they should also be theory based to the extent possible. The question is just to what extent that is. The most recurring and/or intriguing themes from the "stages-of-development" literature should, of course, be put to test. Strong *concepts* from various more fundamental theories in economics and management undoubtedly have their place in a process context as well. However, most established theories arguably remain relatively silent on the process issues themselves, that is, on how the realities

represented by those concepts interrelate and develop over time. Therefore, the topic of growth processes is arguably an area where some exploration is not only excusable, but badly needed.

It was argued above that the influence of antecedents on the amount of growth was the most thoroughly researched area of those implied by Figure 13-1. Our review also showed that quite a number of broad generalizations can be made regarding such relationships. Nevertheless, it can also be argued that this is not the type of results that best serves the needs of management practice. This is so in part because these “growth factors” are often variables that the manager can do little about, and in part because the relationships represent probabilistic truths that may not bear much truth at all in most individual cases. That is, the most relevant “growth factor” in each individual case may be some idiosyncratic factor that is not even represented by the generic variables used in research, or at least a much more concrete manifestation of such a factor on which the research naturally stays silent. Thus, it can be questioned whether broadly based generalizations about the antecedents of growth can *ever* be precise enough to be of much immediate value for management practitioners.

It may be speculated that there should actually be more communality across firms as regards what management challenges different forms of growth *leads* to, regardless of what “success factors” first led to that growth. If so, the type of study that holds most promise from the perspective of furthering management practice (and, hence, education) would be one that combines aspects of *Amount*, *Mode* and *Process*, and related them to *Effects* in terms of a range of management challenges such as acquiring and coordinating a growing resource base, adapting organizational structures and systems, and effectively dealing with recruiting, training, promotions and other people issues in the growing firm. This no doubt partly coincides with what the literature on “stages,” “life-cycles” and “management transitions” have tried to address, although these literatures often postulate a singular process and allow for only a very narrow range of growth modes. It also represents an expanded version of the type of process study we advocated above, with all its research challenges—and more. It is a research task that would require a comprehensive, multi-year program under competent and dedicated leadership, but one which—if successfully undertaken—would really make a difference.

Most researchers will, of course, never get the opportunity to design and carry out such an effort. Perhaps they will never to be part of one, either. Fortunately, there is and will be room for more restricted contributions as well, for example on the growth effects of a couple of factors highlighted by some particular theory. In explicitly starting from a well articulated theory such a study would already be an improvement relative to most of the predecessors. As our review has highlighted, the other ways in which there is opportunity for improvement largely concern the classical research virtues of making sure the

sample and the measures match with the theory. It is thanks to previous research that we now can understand how we can do a better job on that. Finally, future studies should either make a strong case for why firm growth is interesting in its own right or explicitly include in the design those outcomes that growth is otherwise only assumed to lead to.

## REFERENCES

- Acs, Z. J. and Audretsch, D. B. (2003). Innovation and Technological Change. In *Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds), pp. 55–79. Kluwer Academic Publishers, Boston.
- Acs, Z. J. and Preston, L. (1997). Small and Medium-Sized Enterprises, Technology, and Globalization: Introduction to a Special Issue on Small and Medium-Sized Enterprises in the Global Economy. *Small Business Economics* 9 (1): 1–6.
- Adizes, I. (1989). *Corporate Lifecycles: How and Why Corporations Grow and Die and What to Do About It*. Prentice Hall, Englewood Cliffs, NJ.
- Aldrich, H. E. (1999). *Organizations Evolving*. Sage Publications, Newbury Park, CA.
- Aldrich, H. E. and Auster, E. R. (1986). Even Dwarfs Started Small: Liabilities of Age and Size and Their Strategic Implications. In *Research in Organizational Behavior*, Volume 8, B. M. Staw and L. L. Cummings (Eds), pp. 165–190. JAI Press, Greenwich, CT.
- Arbaugh, J. B. and Camp, S. M. (2000). Managing Growth Transitions: Theoretical Perspectives and Research Directions. In *The Blackwell Handbook of Entrepreneurship*, D. L. Sexton and H. Landström (Eds), pp. 308–328. Blackwell, Oxford.
- Ardishvili, A., Cardozo, S., Harmon, S. and Vadakath, S. (1998). *Towards a Theory of New Venture Growth*. Paper presented at the Babson Entrepreneurship Research Conference, Ghent, Belgium.
- Arrow, K. (1983). Innovation in Small and Large Firms. In *Entrepreneurship*, J. Ronen (Ed.), pp. 15–28. Lexington Books, Lexington, MA.
- Audretsch, D. (1995). Innovation, Growth and Survival. *International Journal of Industrial Organization* 13: 441–457.
- Autio, E., Sapienza, H. J. and Almeida, J. G. (2000). Effects of Age at Entry, Knowledge Intensity, and Imitability on International Growth. *Academy of Management Journal* 43: 909–924.
- Bamford, C. E., Dean, T. J. and McDougall, P. P. (1997). Initial Strategies and New Venture Growth: An Examination of the Effectiveness of Broad vs. Narrow Breadth Strategies. In *Frontiers of Entrepreneurship Research 1997*, P. D. Reynolds, W. D. Bygrave, N. M. Carter, P. Davidsson, W. B. Gartner, C. M. Mason and P. P. McDougall (Eds), pp. 375–389. Babson College, Wellesley, MA.
- Barber, J., Metcalfe, J. S. and Porteous, M. (1989). *Barriers to Growth in Small Firms*. Routledge, London.
- Barker, R. G. and Gump, P. V. (1964). *Big School, Small School*. Stanford University Press, Stanford, CA.
- Barkham, R., Gudgin, G., Hart, M. and Hanvey, E. (1996). *The Determinants of Small Firm Growth*, Volume 12. Athenaemum Press, Gateshead, U.K.
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management* 17 (1): 99–120.
- Barney, J. B. (1997). *Gaining and Sustaining Competitive Advantage*. Addison-Wesley, Menlo Park, CA.

- Baum, J. R. and Locke, E. A. (2004). The Relationship of Entrepreneurial Traits, Skill, and Motivation to Subsequent Venture Growth. *Journal of Applied Psychology* 89 (4): 587–598.
- Baum, J. R., Locke, E. A. and Smith, K. G. (2001). A Multidimensional Model of Venture Growth. *Academy of Management Journal* 44 (2): 292–303.
- Baum, J. R., Locke, E. A. and Kirkpatrick, S. A. (1998). A Longitudinal Study of the Relation of Vision and Vision Communication to Venture Growth and Performance. *Journal of Applied Psychology* 83 (1): 43–54.
- Becchetti, L. and Trovato, G. (2002). The Determinants of Firm Growth for Small and Medium Sized Firms. The Role of the Availability of External Finance. *Small Business Economics* 19 (4): 291–306.
- Benito, G. R. G. and Welch, L. S. (1997). De-internationalization. *Management International Review* 37 (2): 7–25.
- Birch, D. L. (1987). *Job Creation in America: How the Smallest Companies Put the Most People to Work*. The Free Press, New York.
- Birch, D. L. and Medoff, J. (1994). Gazelles. In *Labor Markets, Employment Policy and Job Creation*, L. C. Solmon and A. R. Levenson (Eds). Westview Press, Boulder, CO.
- Birch, D. L., Haggerty, A. and Parsons, W. (1995). *Who's Creating Jobs*. Cognetics, Inc, Boston, MA.
- Birley, S. (1985). The Role of Networks in the Entrepreneurial Process. *Journal of Business Venturing* 3 (1): 107–117.
- Birley, S. and Stockley, S. (2000). Entrepreneurial Teams and Venture Growth. In *The Blackwell Handbook of Entrepreneurship*, D. L. Sexton and H. Landström (Eds), pp. 287–307. Blackwell, Oxford.
- Bloodgood, J. M. (1996). The Internationalization of High-Potential U.S. Ventures: Antecedents and Outcomes. *Entrepreneurship Theory and Practice* 8 (4): 61–77.
- Bolton, J. E. (1971). *Small Firms. Report of the Committee of Inquiry on Small Firms*. Her Majesty's Stationery Office, London.
- Butler, J. E. and Hansen, G. S. (1991). Network Evolution, Entrepreneurial Success, and Regional Development. *Entrepreneurship and Regional Development* 3 (1): 1–16.
- Calof, J. and Beamish, P. (1995). Adapting to Foreign Markets: Explaining Internationalization. *International Business Review* 4 (2): 115–131.
- Capon, N., Farley, J. U. and Hoenig, S. (1990). Determinants of Financial Performance: A Meta-Analysis. *Management Science* 36 (10): 1143–1159.
- Carlsson, B. (2002). Institutions, Entrepreneurship, and Growth: Biomedicine and Polymers in Sweden and Ohio. *Small Business Economics* 19 (2): 105–121.
- Carroll, G. R. and Delacroix, J. (1982). Organizational Mortality in the Newspaper Industries of Argentina and Ireland: An Ecological Approach. *Administrative Science Quarterly* 27 (2): 169–198.
- Carroll, G. R. and Hannan, M. T. (2000). *The Demography of Corporations and Industries*. Princeton University Press, Princeton, NJ.
- Casson, M. C. (2000). An Entrepreneurial Theory of the Firm. In *Competence, Governance, and Entrepreneurship*, N. J. Foss and V. Mahnke (Eds), pp. 116–145. Oxford University Press, Oxford.
- Chandler, G. N. and Hanks, S. H. (1994). Founder Competence, the Environment, and Venture Performance. *Entrepreneurship Theory and Practice* 18 (3): 77–89.
- Chandler, G. N. and Jansen, E. (1992). The Founder's Self-Assessed Competence and Venture Performance. *Journal of Business Venturing* 7 (3): 223–236.
- Chandler, G. N., McKelvie, A. and Davidsson, P. (2005). *New Venture Growth: A Transaction Cost Perspective*. Paper presented at the Academy of Management Conference, Honolulu.

- Chetty, S. (1999). Dimensions of Internationalization of Manufacturing Firms in the Apparel Industry. *European Journal of Marketing* 33 (1–2): 121–142.
- Chetty, S. and Campbell-Hunt, C. (2003). Explosive International Growth and Problems of Success Amongst Small to Medium-Sized Firms. *International Small Business Journal* 21 (1): 5–27.
- Cliff, J. E. (1998). Does One Size Fit All? Exploring the Relationship Between Attitudes Towards Growth, Gender, and Business Size. *Journal of Business Venturing* 13 (6): 523–542.
- Cooper, A. C., Gimeno-Gascon, F. J. and Woo, C. Y. (1994). Initial Human and Financial Capital as Predictors of New Venture Performance. *Journal of Business Venturing* 9 (5): 371–395.
- Covin, J. G. and Slevin, D. P. (1997). High Growth Transitions: Theoretical Perspectives and Suggested Directions. In *Entrepreneurship 2000*, D. L. Sexton and R. W. Smilor (Eds), pp. 99–126. Upstart, Chicago.
- Cowling, M. (2004). The Growth-Profit Nexus. *Small Business Economics* 22 (1): 1–9.
- Cox, L. W., Camp, S. M. and Ensley, M. D. (2002). *Does it Pay to Grow? The Impact of Growth on Profitability and Wealth Creation*. Paper presented at the Babson/Kauffman Foundation Entrepreneurship Research Conference, Boulder, CO.
- Cressy, R. and Olofsson, C. (1996). The Financial Conditions for Swedish SMEs: Survey and Research Agenda. *Small Business Economics* 9 (2): 179–194.
- Davidsson, P. (1989a). *Continued Entrepreneurship and Small Firm Growth*. Doctoral Dissertation. Stockholm School of Economics, Stockholm.
- Davidsson, P. (1989b). Entrepreneurship—and After? A Study of Growth Willingness in Small Firms. *Journal of Business Venturing* 4 (3): 211–226.
- Davidsson, P. (1991). Continued Entrepreneurship: Ability, Need, and Opportunity as Determinants of Small Firm Growth. *Journal of Business Venturing* 6 (6): 405–429.
- Davidsson, P. (2003). The Domain of Entrepreneurship Research: Some Suggestions. In *Cognitive Approaches to Entrepreneurship Research*, Volume 6, J. Katz and D. Shepherd (Eds), pp. 315–372. Elsevier/JAI Press, Oxford.
- Davidsson, P. (2004). *Researching Entrepreneurship*. Springer, New York.
- Davidsson, P. and Delmar, F. (1997). *High-Growth Firms and Their Contribution to Employment: The Case of Sweden 1987–96*. OECD Working Party on SMEs, Paris.
- Davidsson, P. and Delmar, F. (1998). Some Important Observations Concerning Job Creation by Firm Size and Age. In *Renaissance der KMU in einer globalisierten Wirtschaft*, H. J. Pleitner (Ed.), pp. 57–67. KMU Vlg HSG, St. Gallen, Switzerland.
- Davidsson, P. and Delmar, F. (2001). Les Entreprises à Forte Croissance et leur Contribution à L'emploi: Le Cas de la Suède 1987–1996. *Revue Internationale PME* 14 (3–4): 164–187.
- Davidsson, P. and Delmar, F. (2003). Hunting for New Employment: The Role of High-Growth Firms. In *Small Firms and Economic Development in Developed and Transition Economies: A Reader*, D. Kirby and A. Watson (Eds), pp. 7–20. Ashgate, Aldershot, U.K.
- Davidsson, P. and Henrekson, M. (2002). Institutional Determinants of the Prevalence of Start-ups and High-Growth Firms: Evidence from Sweden. *Small Business Economics* 19 (2): 81–104.
- Davidsson, P. and Wiklund, J. (2000). Conceptual and Empirical Challenges in the Study of Firm Growth. In *The Blackwell Handbook of Entrepreneurship*, D. Sexton and H. Landström (Eds), pp. 26–44. Blackwell Business, Oxford, MA.
- Davidsson, P., Delmar, F. and Wiklund, J. (2002). Entrepreneurship as Growth; Growth as Entrepreneurship. In *Strategic Entrepreneurship: Creating a New Mindset*, M. A. Hitt, R. D. Ireland, S. M. Camp and D. L. Sexton (Eds), pp. 328–342. Blackwell, Oxford.
- Davidsson, P., Lindmark, L. and Olofsson, C. (1996). *Näringslivsdynamik under 90-talet (Business Dynamics in the 90s)*. NUTEK, Stockholm.
- Davidsson, P., Lindmark, L. and Olofsson, C. (1998). Smallness, Newness and Regional Development. *Swedish Journal of Agricultural Research* 28 (1): 57–71.

- Davidsson, P., Steffens, P. and Fitzsimmons, J. (2005). *Growing Profitable or Growing from Profits: Putting the Horse in Front of the Cart?* Paper presented at the Academy of Management Meeting, Honolulu.
- Davis, S. J., Haltiwanger, J. and Schuh, S. (1996a). *Job Creation and Destruction*. MIT Press, Boston, MA.
- Davis, S. J., Haltiwanger, J. and Schuh, S. (1996b). Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts. *Small Business Economics* 8 (4): 297–315.
- Delmar, F. (1997). Measuring Growth: Methodological Considerations and Empirical Results. In *Entrepreneurship and SME Research: On its Way to the Next Millennium*, R. Donckels and A. Miettinen (Eds), pp. 190–216. Ashgate, Aldershot, U.K.
- Delmar, F. and Davidsson, P. (1999). Firm Size Expectations of Nascent Entrepreneurs. In *Frontiers of Entrepreneurship Research 1999*, P. D. Reynolds, W. D. Bygrave, S. Manigart, C. Mason, G. D. Meyer, H. J. Sapienza and K. G. Shaver (Eds), pp. 90–104. Babson College, Wellesley, MA.
- Delmar, F. and Wiklund, J. (2003). *The Effect of the Entrepreneur's Growth Motivation on Subsequent Growth: A Longitudinal Study*. Paper presented at the Academy of Management Meeting, Seattle.
- Delmar, F., Davidsson, P. and Gartner, W. (2003). Arriving at the High-Growth Firm. *Journal of Business Venturing* 18 (2): 189–216.
- Dennis, W. J. and Solomon, G. (2001). *Changes in Intention to Grow over Time*. Paper presented at the Babson/Kauffman Foundation Entrepreneurship Research Conference, Jönköping, Sweden.
- Dess, G. G. and Beard, D. W. (1984). Dimensions of Organizational Task Environments. *Administrative Science Quarterly* 29 (1): 52–73.
- Donckels, R. and Lambrecht, J. (1995). Networks and Small Business Growth: An Explanatory Model. *Small Business Economics* 7 (4): 273–289.
- DuRietz, A. and Henrekson, M. (2000). Testing the Female Underperformance Hypothesis. *Small Business Economics* 14 (1): 1–10.
- Eisenhardt, K. M. and Schoonhoven, C. B. (1990). Organizational Growth: Linking Founding Team, Strategy, Environment and Growth among U.S. Semiconductor Ventures, 1978–1988. *Administrative Science Quarterly* 35 (3): 504–530.
- Ensley, M. D., Pearson, A. W. and Amason, A. S. (2002). Understanding the Dynamics of New Venture Top Management Teams: Cohesion, Conflict and New Venture Performance. *Journal of Business Venturing* 17 (4): 365–386.
- Fischer, E., Reuber, R. and Dyke, L. (1993). A Theoretical Overview and Extension of Research on Sex, Gender, and Entrepreneurship. *Journal of Business Venturing* 8 (2): 151–168.
- Flamholtz, E. G. (1986). *Managing the Transition from an Entrepreneurship to a Professionally Managed Firm*. Jossey-Bass, San Francisco.
- Flamholtz, E. G. and Randle, Y. (1990). *Growing Pains: How to Make the Transition from Entrepreneurship to a Professionally Managed Firm*. Jossey-Bass, San Francisco.
- Fombrun, C. J. and Wally, S. (1989). Structuring Small Firms for Rapid Growth. *Journal of Business Venturing* 4 (2): 107–222.
- Freeman, J., Carroll, G. R. and Hannan, M. T. (1983). The Liability of Newness: Age Dependence in Organizational Death Rates. *American Sociological Review* 48 (5): 692–710.
- Galbraith, J. (1982). The Stages of Growth. *Journal of Business Strategy* 3 (1): 70–79.
- Garney, E. (1998). A Theory of the Early Growth of the Firm. *Industrial and Corporate Change* 7 (3): 523–556.
- Gartner, W. B. (1988). 'Who is an Entrepreneur' is the Wrong Question. *American Small Business Journal* 12 (4): 11–31.

- Gartner, W. B. and Carter, N. (2003). Entrepreneurial Behavior and Firm Organizing Processes. In *Handbook of Entrepreneurship Research*, Z. J. Acs and D. B. Audretsch (Eds), pp. 195–221. Kluwer, Dordrecht.
- Gibb, A. A. and Davies, L. G. (1990). In Pursuit of Frameworks for the Development of Growth Models of the Small Business. *International Small Business Journal* 9 (1): 15–31.
- Gray, C. (1990). *Entrepreneurial Motivation and the Smaller Business*. Paper presented at the 15th IAREP Colloquium, Exeter, U.K.
- Greening, D. and Barringer, B. (1996). A Qualitative Study of Managerial Challenges Facing Small Business Geographic Expansion. *Journal of Business Venturing* 11 (4): 233–256.
- Greiner, L. E. (1972). Evolutions and Revolutions as Organizations Grow. *Harvard Business Review* 50 (4): 37–46.
- Hambrick, D. C. and Crozier, L. M. (1985). Stumblers and Stars in the Management of Rapid Growth. *Journal of Business Venturing* 1 (1): 31–45.
- Hanks, S. H., Watson, C. J., Jansen, E. and Chandler, G. N. (1993). Tightening the Life-cycle Construct: A Study of Growth Stage Configurations in High-Technology Organizations. *Entrepreneurship Theory and Practice* 18 (2): 5–29.
- Hannan, M. T. and Freeman, J. H. (1977). The Population Ecology of Organizations. *American Journal of Sociology* 82 (5): 929–964.
- Hansen, E. L. (1995). Entrepreneurial Networks and New Organization Growth. *Entrepreneurship Theory and Practice* 19 (4): 7–19.
- Hofer, C. W. and Charan, R. (1984). The Transition to Professional Management: Mission Impossible? *American Journal of Small Business* 9 (1): 1–11.
- Hoy, F., McDougall, P. P. and Dsouza, D. E. (1992). Strategies and Environments of High Growth Firms. In *The State of the Art of Entrepreneurship*, D. L. Sexton and J. D. Kasarda (Eds), pp. 341–357. PWS-Kent Publishing, Boston, MA.
- Human, S. E. and Matthews, C. H. (2004). Future Expectations for the New Business. In *Handbook of Entrepreneurial Dynamics: the Process of Business Creation*, W. B. Gartner, K. G. Shaver, N. M. Carter and P. D. Reynolds (Eds), pp. 94–103. Sage, Thousand Oaks, CA.
- Ibeh, N. (2003). Toward a Contingency Framework of Export Entrepreneurship: Conceptualizations and Empirical Evidence. *Small Business Economics* 20 (1): 49–58.
- Jarillo, J. C. (1989). Entrepreneurship and Growth: The Strategic Use of External Resources. *Journal of Business Venturing* 4 (2): 133–147.
- Jenkins, M. and Johnson, G. (1997). Entrepreneurial Intentions and Outcomes: A Comparative Causal Mapping Study. *Journal of Management Studies* 34 (6): 895–920.
- Johanson, J. and Vahlne, J.-E. (1977). The Internationalization Process of the Firm: A Model of Knowledge Development and Increasing Foreign Market Commitments. *Journal of International Business Studies* 8 (1): 23–32.
- Johanson, J. and Wiedersheim-Paul, F. (1975). Internationalization of the Firm: Four Swedish Cases. *Journal of Management Studies* 12 (3): 305–322.
- Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica* 50 (3): 649–670.
- Julien, P.-A. and Ramangalahy, C. (2003). Competitive Strategy and Performance of Exporting SMEs: An Empirical Investigation of the Impact of Their Export Information Search and Competences. *Entrepreneurship Theory and Practice* 27 (3): 227–245.
- Kazanjian, R. K. (1988). Relation of Dominant Problems to Stages of Growth in Technology-Based New Ventures. *Academy of Management Journal* 31 (2): 257–279.
- Kazanjian, R. K. and Drazin, R. (1989). An Empirical Test of Stage of Growth Progression Model. *Management Science* 35 (12): 1489–1503.
- Killing, J. P. (1978). Diversification through Licensing. *R&D Management* 8 (3): 159–163.
- Kish, L. (1987). *Statistical Design for Research*. Wiley, New York.

- Kolvereid, L. and Bullvåg, E. (1996). Growth Intentions and Actual Growth: The Impact of Entrepreneurial Choice. *Journal of Enterprising Culture* 4 (1): 1–17.
- Kraemer, T. D. and Venkataraman, S. (1997). Extraordinary Feats of Entrepreneurial Enterprise: Strategies for Rapid, Sustained Growth. In *Entrepreneurship in a Global Context*, S. Birley and I. C. MacMillan (Eds), pp. 82–107. Routledge, London.
- Levie, J. (1997). Patterns of Growth and Performance: An Empirical Study of Young, Growing Ventures in France, Ireland and Scotland. In *Frontiers of Entrepreneurship Research 1997*, P. D. Reynolds, W. D. Bygrave, N. M. Carter, P. Davidsson, W. B. Gartner, C. M. Mason and P. P. McDougall (Eds), pp. 375–389. Babson College, Wellesley, MA.
- Low, M. B. and MacMillan, I. C. (1988). Entrepreneurship: Past Research and Future Challenges. *Journal of Management* 14 (2): 139–161.
- Lumpkin, G. T. and Dess, G. G. (2001). Linking Two Dimensions of Entrepreneurial Orientation to Firm Performance: The Moderating Role of Environment and Industry Life Cycle. *Journal of Business Venturing* 16 (5): 429–451.
- Lumpkin, G. T. and Dess, G. G. (1996). Clarifying the Entrepreneurial Orientation Construct and Linking It to Performance. *Academy of Management Review* 21 (1): 135–172.
- Madsen, T. K. and Servais, P. (1997). The Internationalization of Born Globals: An Evolutionary Process? *International Business Review* 6 (3): 82–93.
- Markman, G. D. and Gartner, W. B. (2002). Is Extraordinary Growth Profitable? A Study of *Inc.* 500 High-Growth Companies. *Entrepreneurship Theory and Practice* 27 (1): 65–75.
- McCann, J. (1991). Patterns of Growth, Competitive Technology, and Financial Strategies in Young Ventures. *Journal of Business Venturing* 6 (3): 198–208.
- McDougall, P. P. and Oviatt, B. M. (1997). International Entrepreneurship Literature in 1990s and Directions for Future Research. In *Entrepreneurship 2000*, D. L. Sexton and R. W. Smilor (Eds), pp. 291–320. Upstart, Chicago, IL.
- McDougall, P., Covin, J., Robinson, R. and Herron, L. (1994). The Effect of Industry Growth and Strategic Breadth on New Venture Performance and Strategy Content. *Strategic Management Journal* 15 (7): 537–554.
- McDougall, P. P., Shane, S. and Oviatt, B. M. (1994). Explaining the Formation of International New Ventures: The Limits of Theories from International Business Research. *Journal of Business Venturing* 9 (5): 469–487.
- Melin, L. (1992). Internationalization as Strategy Process. *Strategic Management Journal* 13: 161–179.
- Mendelson, H. (2000). Organizational Architecture and Success in the Information Technology Industry. *Management Science* 46 (4): 513–529.
- Meyer, G. D., Neck, H. M. and Meeks, M. D. (2002). The Entrepreneurship-Strategic Management Interface. In *Strategic Entrepreneurship: Creating a New Mindset*, M. A. Hitt, R. D. Ireland, M. S. Camp and D. S. Sexton (Eds), pp. 19–44. Blackwell, Oxford.
- Miller, D. and Friesen, P. H. (1984). A Longitudinal Study of the Corporate Life-Cycle. *Management Science* 30 (10): 1161–1183.
- Mok, A. L. and van den Tillaart, H. (1990). Farmers and Small Businessmen: A Comparative Analysis of their Careers and Occupational Orientation. In *New Findings and Perspectives in Entrepreneurship*, R. Donckels and A. Miettinen (Eds), pp. 203–230. Avebury, Aldershot.
- Mosakowski, E. (2002). Overcoming Resource Disadvantages in Entrepreneurial Firms: When Less is More. In *Strategic Entrepreneurship: Creating a New Mindset*, M. A. Hitt, R. D. Ireland, M. S. Camp and D. S. Sexton (Eds), pp. 106–126. Blackwell, Oxford.
- Neumair, U. (1998). *A General Model of Corporate Failure and Survival: A Complexity Theory Approach*. Doctoral Dissertation. University of St. Gallen, St. Gallen, Switzerland.
- O’Gorman, C. (1997). Success Strategies in High Growth Small and Medium-Sized Enterprises. In *Technology, Innovation and Enterprise*, D. Jones-Evans and M. Klofsten (Eds), pp. 179–208. MacMillan, London.



- O'Farrell, P. N. and Hitchens, D. M. W. N. (1988). Alternative Theories of Small-Firm Growth: A Critical Review. *Environment and Planning A* 20 (2): 1365–1383.
- O'Rand, A. M. and Krecker, M. L. (1990). Concepts of the Life-Cycle: Their History, Meanings, and Uses in the Social Sciences. *Annual Review of Sociology* 16 (1): 241–262.
- Penrose, E. (1959). *The Theory of the Growth of the Firm*. Oxford University Press, Oxford.
- Porter, M. E. (1980). *Competitive Strategy*. Free Press, New York.
- Quinn, R. E. and Cameron, K. S. (1983). Organizational Life Cycles and Shifting Criteria of Effectiveness: Some Preliminary Evidence. *Management Science* 29 (1): 33–51.
- Raffa, M., Zollo, G. and Caponi, R. (1996). The Development Process of Small Firms. *Entrepreneurship and Regional Development* 8 (4): 359–372.
- Reynolds, P. D. and White, S. B. (1997). *The Entrepreneurial Process: Economic Growth, Men, Women, and Minorities*. Quorum Books, Westport, CT.
- Reynolds, P. D., Bygrave, W. D. and Autio, E. (2003). *GEM 2003 Global Report*. Kauffman Foundation, Kansas.
- Riding, A. L. and Haines, G. J. (1998). Defaulting on Loan Guarantees: Costs and Benefits of Encouraging Early-Stage Growth. In *Frontiers of Entrepreneurship Research 1998*, W. D. Bygrave, N. M. Carter, S. Manigart, G. D. Meyer and K. G. Shaver (Eds), pp. 504–518. Babson College, Wellesley, MA.
- Roberts, E. B. and Berry, C. A. (1985). Entering New Businesses: Selecting Strategies for Success. *Sloan Management Review* 26: 3–17.
- Robson, P. J. A. and Bennett, R. J. (2000). SME Growth: The Relationship with Business Advice and External Collaboration. *Small Business Economics* 15 (3): 193–208.
- Rollag, K. (2001). How Fast Growth Promotes Rapid Socialization in Entrepreneurial Firms. In *Frontiers of Entrepreneurship Research 2001*, W. D. Bygrave, E. Autio, C. G. Brush, P. Davidsson, P. G. Green, P. D. Reynolds and H. J. Sapienza (Eds). Babson College, Wellesley, MA.
- Roper, S. (1999). Modeling Small Business Growth and Profitability. *Small Business Economics* 13 (3): 235–252.
- Ruef, M., Aldrich, H. E. and Carter, N. M. (2003). The Structure of Organizational Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. *American Sociological Review* 68 (2): 195–222.
- Samuelsson, M. (2004). *Creating New Ventures: A Longitudinal Investigation of the Nascent Venturing Process*. Doctoral Dissertation, Jönköping International Business School, Jönköping.
- Samuelsson, M. (2001). Modeling the Nascent Venture Opportunity Exploitation Process across Time. In *Frontiers of Entrepreneurship Research 2001*, W. D. Bygrave, E. Autio, C. G. Brush, P. Davidsson, P. G. Green, P. D. Reynolds and H. J. Sapienza (Eds). Babson College, Wellesley, MA.
- Sandberg, W. R. and Hofer, C. W. (1987). Improving New Venture Performance: The Role of Strategy, Industry Structure, and the Entrepreneur. *Journal of Business Venturing* 2 (2): 5–28.
- Sapienza, H. J., Korsgaard, M. A. and Forbes, D. P. (2003). The Self-Determination Motive and Entrepreneurs' Choice of Financing. In *Cognitive Approaches to Entrepreneurship Research*, Volume 6, J. Katz and D. Shepherd (Eds), pp. 107–140. Elsevier/JAI Press, Oxford.
- Scott, M. and Bruce, R. (1987). Five Stages of Growth in Small Business. *Long Range Planning* 20 (3): 45–52.
- Sexton, D. L. (1997). Entrepreneurship Research Needs and Issues. In *Entrepreneurship 2000*, D. L. Sexton and R. W. Smilor (Eds), pp. 401–408. Upstart Publishing Company, Chicago, IL.
- Sexton, D. L., Pricer, R. W. and Nenide, B. (2000). *Measuring Performance in High-Growth Firms*. Paper presented at the Babson/Kauffman Foundation Entrepreneurship Research Conference, Babson College, MA.

- Shane, S. and Venkataraman, S. (2000). The Promise of Entrepreneurship as a Field of Research. *Academy of Management Review* 25 (1): 217–226.
- Stanworth, J. and Curran, J. (1973). *Management Motivation in the Smaller Business*. Gower Press, Epping, U.K.
- Stevenson, H. H. and Jarillo, J. C. (1990). A Paradigm of Entrepreneurship: Entrepreneurial Management. *Strategic Management Journal* 11: 17–27.
- Stinchcombe, A. L. (1965). Social Structure and Organizations. In *Handbook of Organizations*, J. D. March (Ed.), pp. 142–193. Rand McNally, Chicago, IL.
- Storey, D. J. (1997). *The Ten Percenters. Second Report*. Deloitte & Touche, London.
- Storey, D. J. (1994). *Understanding the Small Business Sector*. Routledge, London.
- Stubbart, C. I. and Smalley, R. D. (1999). The Deceptive Allure of Stage Models of Strategic Processes. *Journal of Management Inquiry* 8 (3): 273–286.
- Tang, N. K. H., Jones, O. and Forrester, P. L. (1997). Organizational Growth Demands Concurrent Engineering. *Integrated Manufacturing Systems* 8 (1): 29–34.
- Thorelli, H. B. (1987). Entrepreneurship in International Marketing: Some Research Opportunities. In *Research at the Marketing-Entrepreneurship Interface*, G. E. Hills (Ed.), pp. 183–204. USASBE, Marietta, GA.
- Watson, J. (2002). Comparing the Performance of Male- and Female-Controlled Business: Relating Outputs to Inputs. *Entrepreneurship: Theory & Practice* 26 (3): 91–100.
- Weinzimmer, L. G., Nystrom, P. C. and Freeman, S. J. (1998). Measuring Organizational Growth: Issues, Consequences and Guidelines. *Journal of Management* 24 (2): 235–262.
- Welch, D. and Welch, L. S. (1996). The Internationalization Process and Networks: A Strategic Management Perspective. *Journal of International Marketing* 4 (3): 11–27.
- Wernerfelt, B. (1995). The Resource-Based View of the Firm: Ten Years After. *Strategic Management Journal* 16 (3): 171–174.
- Wernerfelt, B. (1984). A Resource Based View of the Firm. *Strategic Management Journal* 5 (2): 171–180.
- Whetten, D. A. (1987). Organizational Growth and Decline Processes. *Annual Review of Sociology* 13: 335–358.
- Wiklund, J. (1998). *Small Firm Growth and Performance: Entrepreneurship and Beyond*. Doctoral Dissertation. Jönköping International Business School, Jönköping.
- Wiklund, J. (1999). The Sustainability of the Entrepreneurial Orientation-Performance Relationship. *Entrepreneurship Theory & Practice* 24 (1): 37–48.
- Wiklund, J. (2001). Growth Motivation and Its Influence on Subsequent Growth. In *Frontiers of Entrepreneurship Research 2001*, W. D. Bygrave, E. Autio, C. G. Brush, P. Davidsson, P. G. Green, P. D. Reynolds and H. J. Sapienza (Eds). Babson College, Wellesley, MA.
- Wiklund, J. (2005). *Knowledge Accumulation In Growth Studies: The Consequences of Methodological Choices*. Paper presented at ERIM Expert Workshop ‘Perspectives on the Longitudinal Analysis of New Firm Growth’, Erasmus University, Rotterdam.
- Wiklund, J. and Davidsson, P. (1999). *A Resource-based View on Organic and Acquired Growth*. Paper presented at the Academy of Management Conference, Chicago.
- Wiklund, J. and Shepherd, D. (2003). Aspiring For, and Achieving Growth: The Moderating Role of Resources and Opportunities. *Journal of Management Studies* 40 (8): 1911–1941.
- Wiklund, J. and Shepherd, D. (2005). Entrepreneurial Orientation and Small Business Performance: A Configurational Approach. *Journal of Business Venturing* 20 (1): 71–91.
- Wiklund, J., Davidsson, P. and Delmar, F. (2003). Expected Consequences of Growth and Their Effect on Growth Willingness in Different Samples of Small Firms. *Entrepreneurship Theory & Practice* 27 (3): 247–269.
- Winborg, J. and Landström, H. (2001). Financial Bootstrapping in Small Businesses: Examining Small Business Managers’ Resource Acquisition Behaviors. *Journal of Business Venturing* 16 (3): 235–254.

- Woo, C. Y., Cooper, A. C. and Dunkelberg, W. C. (1991). The Development and Interpretation of Entrepreneurial Typologies. *Journal of Business Venturing* 6 (2): 93–114.
- Zahra, S. (1991). Predictors and Financial Outcomes of Corporate Entrepreneurship: An Exploratory Study. *Journal of Business Venturing* 6 (4): 259–285.

## Stage 5: Venture Development II: Social Issues

HELEN HAUGH

*Judge Business School, University of Cambridge, Trumpington Street, Cambridge CB2 1AG,  
United Kingdom*

*E-mail: h.haugh@jbs.cam.ac.uk*

## 14. Nonprofit Social Entrepreneurship

### 1. INTRODUCTION

Oxfam, the YMCA and Emmaus are nonprofit organizations that pursue entrepreneurial strategies to create social value. Their engagement with revenue generation distinguishes them from traditional nonprofit organizations as much as their nonprofit status differentiates them from mainstream commercial enterprises.

The term “social enterprise” describes organizations, such as these, that aim to achieve financial sustainability through trading for a social purpose. Any surplus they make is usually re-invested to further the approved mission of the organization, and not distributed for private gain. They are located mainly in the nonprofit sector, although several, such as The Big Issue and the Grameen Bank, have been incorporated as for-profit social enterprises. This chapter focuses on nonprofit social entrepreneurship.

Organizations that are neither publicly nor privately owned (Gui, 1991) and that operate to promote economic and social well being (OECD, 1999) are collectively termed the nonprofit sector. The sector is diverse, extending from large, charitable organizations, such as the Red Cross, to small, informal groups and voluntary associations operating at community level. The organizations within it have in common the pursuit of social and/or environmental goals, either for mutual or societal benefit.

The nonprofit sector is dominated by services (Francois, 2003) and includes a range of legal status and institutional formats (charity, Industrial and Provident Society, mutual organization, trust, company limited by guarantee, etc.); economic sectors (such as housing, education, sport, recreation, arts, and health and social care); mutual organizations (such as credit unions and

cooperatives); and ventures to provide public goods for societal benefit (such as consumer protection and environmental sustainability).

Globally, the nonprofit sector is an important employer, purchaser, innovator and voice for civil society. It also provides an innovative route for tackling inequality and poverty. Most countries can identify some level of nonprofit activity, and the rise in the number of nonprofit organizations (Frumkin, 2002, p. 141) reflects the increased importance of the sector (Salamon and Anheier, 1997, p. 1). At country level, nonprofit organizations have emerged in response to legal, political, economic, social and cultural forces (Holtmann, 1988; DiMaggio and Anheier, 1990; Ben-Ner and Van Hoomissen, 1991).

The nonprofit sector is characterized as being dynamic and responsive, with many organizations adopting entrepreneurial strategies and seeking financial sustainability through trading (Weisbrod, 2004). This reflects trends in: government policies (the spreading of democracy, Third Way politics and the changing role of government from delivery agent to contractor); economics (persistent and systemic inequality and poverty); and society (the rise of civil society, the decline in social capital (Putnam, 2000), and an upsurge in organized, voluntary activity (Salamon et al., 2003, p. 1)). It is also a response to greater pressure on financial resources due to increased competition for decreasing donations and grants (Dees, 1998) as well as the potential application of information and communication technologies to promote economic and social inclusion.

This chapter begins by summarizing the size, scale and defining characteristics of the nonprofit sector. The rationale for the sector is explained in terms of supply and demand factors. Nonprofit social entrepreneurship is addressed by considering the increasing pressure on nonprofit organizations to adopt more enterprising strategies. The characteristics and motivations of the social entrepreneur and the process of social entrepreneurship are then examined. The process of social venture creation (opportunity recognition, innovation, resource acquisition, enterprise creation and development) and performance measurement is summarized and, in conclusion, four potential research themes are identified for scholars. The literature drawn on comes from nonprofit economics, entrepreneurship and social entrepreneurship, including examples of nonprofit social enterprise from the U.S. and U.K.

## 2. SIZE AND SCALE OF THE NONPROFIT SECTOR

Although many countries gather statistics on indigenous nonprofit activity, the data suffer from lack of international comparability due to variations in information sources, methodology and data collection. Globally, the size and scale of the nonprofit sector has been the subject of longitudinal research by the

TABLE 14-1 *Percentage of paid staff and volunteers by field of activity*

| Activity                                       | Paid staff (%) | Volunteers (%) |
|--|----------------|----------------|
| Culture and recreation                         | 13             | 25             |
| Education and research                         | 30             | 8              |
| Health   | 17             | 8              |
| Social services                                | 18             | 27             |
| Environment                                    | 2              | 3              |
| Development and housing                        | 7              | 10             |
| Civic, advocacy and politics                   | 3              | 7              |
| Foundations                                    | 1              | 2              |
| International                                  | 1              | 1              |
| Business and professional associations, unions | 7              | 6              |
| Other (not classified elsewhere)               | 2              | 3              |

Source: Salamon et al., 2003, pp. 10, 24.

Johns Hopkins Comparative Nonprofit Research Project ([www.jhu.edu](http://www.jhu.edu)). This study has gathered comparable data from researchers about size (measured by employment, volunteers and percentage of GDP), activity and funding from 35 developed, developing and transitional countries (see Table 14-1). While the actual number of nonprofit organizations has not been collected (because of measurement difficulties), data about employment, income source and revenue generation give a broad indication of the economic importance of the sector.

By the late 1990s, the nonprofit sector had aggregate expenditure of \$1.3 trillion, representing 5.1% combined Gross Domestic Product (GDP) of the countries surveyed (Salamon et al., 2003). The workforce consisted of 39.5 million full-time equivalent (FTE) individuals, made up of 22.7 million (57%) paid workers and 16.8 (43%) million volunteers. In total, an average of 4.4% of the economically active population is employed in the sector. In relation to the service sector, 66% of employment was in education, health, and social services, with 40% of volunteer time spent in recreation (including sports) and social services (Salamon et al., 1999).

In the U.S., approximately 8.5 million FTE paid employees and 4.9 million volunteers work in the nonprofit sector. The largest industry sector by paid employees was health (46.3%), followed by education (21.5%). The largest sector by volunteers was social services (36.7%), and then health (13.6%) (Salamon et al., 2003). According to data from the National Center for Charitable Statistics, there were approximately 245,000 operating charities in the U.S. by the end of 1996 (Cordes et al., 2004).

In the U.K., approximately 1.4 million FTE paid employees and 1.1 million volunteers work in the nonprofit sector. The largest industry sector by paid employees is education (41.5%), followed by culture (24.5%). The

largest sector by volunteers is culture (31.3%), followed by development (18%) (Salamon et al., 2003). More recent data from the U.K. noted that there were between 500,000 and 700,000 organizations in the charitable and wider not-for-profit sector (Gov, 2002). This included 180,000 to 360,000 community-level organizations and 188,000 registered charities in England and Wales (Gov, 2002). Although most registered charities are relatively small, the sector employed more than 500,000 FTEs, amounting to 2.2% of the U.K. workforce (Jas et al., 2002).

Although cross-national data are not completely compatible, the broad patterns of employment and volunteer involvement are similar (Rose-Ackerman, 1996). There are significant regional variations in industry dominance: welfare (Western Europe); recreation and culture (Central Europe); education (Latin America); and health (other developed countries). The data endorse the importance of nonprofits in service industries (Rose-Ackerman, 1996; Ben-Ner and Van Hoomissen, 1991; Salamon et al., 2003), their over-representation in labor-intensive activities (Zimmerman, 1999:592) and the responsiveness of the nonprofit sector to influences at country level.

### 3. DEFINING THE NONPROFIT SECTOR

When researchers at Johns Hopkins University ([www.jhu.edu](http://www.jhu.edu)) embarked on the global study of the nonprofit sector in 1991, an immediate challenge involved how to define the boundaries of the sector and identify those organizations to be included in the data collection. Due to inter-country differences in the structure, status and purpose of nonprofit organizations, five criteria were used to distinguish nonprofit organizations from public bodies and private enterprises (Salamon and Anheier, 1997). This resulted in a definition of a nonprofit organization as one that is private, aims to achieve a social purpose, does not distribute profits to those with a controlling interest, is self-governing and that people are free to join or support voluntarily. This section discusses these criteria in more detail.

Nonprofit organizations are separate from the private sector and the state (Mertens, 1999). Although many nonprofit organizations receive government support (in the form of grants and service delivery contracts), they are not part of state infrastructure. To distinguish nonprofit organizations from casual, nonpermanent, transient projects, there must be evidence of organizational permanence and regularity, such as details of membership, minutes from regular meetings and/or other records and administration. This still enables relatively permanent informal groups to be included, as well as legally constituted organizations.

Nonprofit organizations operate primarily to create social value and are subject to the nondistribution constraint (NDC). The NDC prohibits them from



distributing profits to those with a controlling interest in the organization, such as directors, managers and trustees (Hansmann, 1980). Nonprofit organizations have in common the principals of enhancing the social fabric of society, taking an ethical approach to doing business and the nondistribution of profits generated from their business activities. Their overall mission is to achieve a beneficial purpose, for example, by contributing to sustainable growth, shared prosperity, and social and economic justice (OECD, 2003). They may achieve this through philanthropy, advocacy and/or by their activities (Hansmann, 1980).

In most countries, in recognition of the social value that they create, nonprofit organizations are granted exemption from certain taxes. In the U.S., for example, to qualify for nonprofit status an organization must meet specific requirements and limitations stipulated by the Inland Revenue Service. Approved organizations are classified as tax-exempt corporations and are exempt from paying taxes that would be levied on a for-profit business (federal income taxes under sections 501–528 of the Internal Revenue Code). They are also exempt from various local and state taxes, and some are eligible to receive tax-deductible gifts and donations ([www.irs.ustreas.gov](http://www.irs.ustreas.gov)).

In the U.K., nonprofit status confers advantages in terms of access to grant income, various tax incentives (such as being eligible to reclaim basic rate income tax on donations from U.K. tax payers), and business rates relief (a minimum reduction of 80% in Uniform Business Rate in 2002) (Jas et al., 2002). Until 2004, the U.K. law on charities restricted the activities of nonprofit organizations to a pre-specified list of charitable purposes in the public benefit and those that engaged in trading were required to establish a separate trading entity for this purpose. This restriction may be lifted in the future (Gov, 2002) which would increase the entrepreneurial development opportunities across the nonprofit sector.

Although nonprofit organizations are allowed to make profits in their line of business, in exchange for the tax benefits conferred upon them, constraints are imposed as to how surplus revenues can be expended. Constitutionally, nonprofit organizations have been distinguished from for-profit organizations by their absence of stock, or other indicator of ownership that conferred on their owners “a simultaneous share in profits and control” (Hansmann, 1980, p. 838). Thus, if a nonprofit organization has generated a surplus, it is “barred from distributing its net earnings, if any, to individuals who exercise control over it, such as members, officers, directors or trustees” (Hansmann, 1980, p. 838). The NDC also operates internally to constrain implicit surplus distribution via above-market remuneration in the form of large salaries and other perks (DiMaggio and Anheier, 1990, p. 138). This applies only when an organization is of sufficient scale to generate large earnings that cannot plausibly be paid out as reasonable salaries (Hansmann, 1980). The NDC thus prevents

individuals from appropriating contributions (Bilodeau and Slivinski, 1998) and from personally profiting at the expense of donors and/or intended beneficiaries (Oberfield and Dees, 1994).

Nonprofit organizations must be self-governing and in control of their own affairs. In most countries, control is exercised by a board of independent, unremunerated trustees who act as stewards and are responsible for the actions of the organization (Taylor et al., 1996). They have the authority to veto managerial decisions and ultimately to close the organization down. Trustees hold regular board meetings, the timing and frequency of which are recorded in the constitutional documents of the organization.

Finally, nonprofit organizations are voluntary in that membership in them, or participation in their activities, is not legally required, enforceable or otherwise compulsory. The decision to participate in the activities of the organization is independently taken by the individual concerned.

In addition to identifying the six criteria of nonprofit organizations, the John Hopkins study created a formal classification scheme of primary activity, identifying 12 fields of nonprofit activity. This enabled data on the size and scale of the nonprofit sector in more than 35 countries to be mapped (see Table 14-2). The result was that comprehensive information on the size, scale and activity of the nonprofit sector, as well as a clear set of defining characteristics for nonprofit organizations, became available, greatly improving knowledge of the sector.

The following section considers the explanations given to account for the presence of a nonprofit sector in an economy.

#### 4. RATIONALE FOR THE NONPROFIT SECTOR

Traditionally, most national economies have been defined in terms of two sectors: a public, state sector and a private, for-profit sector in which entrepreneurial activity is generally situated. In addition, many countries are able to identify some level of nonprofit activity, although it may be peripheral and only make a marginal contribution to economic wealth. Over the years, however, academic and policy interest in nonprofit organizations has been growing. The reasons for this include: the existence of long-term structural unemployment in many countries; the crisis of excessive demands on state welfare expenditure (Defourny, 2001, p. 1); greater interest in associational relationships; increased participation in civil society and national consumer movements; an overall decline in per capita giving (Salaman and Anheier, 1997); and increased competition for donations (Skloot, 1987). These factors have combined to leave some sectors of society served neither by the market nor the state, leading to the emergence of nonprofit organizations in a residual capacity—acting as provider in conditions of market failure to those in need, or as trusted intermediary in conditions of contract failure.

TABLE 14-2 *Paid staff and volunteers by country*

|                | FTE paid employment<br>(000s) | FTE volunteering<br>(000s) |
|----------------|-------------------------------|----------------------------|
| Argentina      | 395.3                         | 264.1                      |
| Australia      | 402.6                         | 177.1                      |
| Austria        | 143.6                         | 40.7                       |
| Belgium        | 357.8                         | 99.1                       |
| Brazil         | 1034.6                        | 139.2                      |
| Colombia       | 286.9                         | 90.8                       |
| Czech Republic | 74.2                          | 40.9                       |
| Egypt          | 611.9                         | 17.3                       |
| Finland        | 62.8                          | 74.8                       |
| France         | 959.8                         | 1021.7                     |
| Germany        | 1440.9                        | 978.1                      |
| Hungary        | 44.9                          | 9.9                        |
| Ireland        | 118.7                         | 31.7                       |
| Israel         | 145.4                         | 31.3                       |
| Italy          | 568.5                         | 381.6                      |
| Japan          | 2140.1                        | 695.1                      |
| Kenya          | 174.9                         | 112.4                      |
| Mexico         | 93.8                          | 47.2                       |
| Morocco        | 74.5                          | 83.4                       |
| Netherlands    | 661.7                         | 390.1                      |
| Norway         | 60.0                          | 103.0                      |
| Pakistan       | 261.8                         | 180.8                      |
| Peru           | 129.8                         | 80.1                       |
| Philippines    | 187.3                         | 330.3                      |
| Poland         | 122.5                         | 32.1                       |
| Romania        | 37.4                          | 46.5                       |
| Slovakia       | 16.2                          | 6.9                        |
| South Africa   | 298.2                         | 264.3                      |
| South Korea    | 413.3                         | 122.1                      |
| Spain          | 475.2                         | 253.6                      |
| Sweden         | 82.6                          | 260.3                      |
| Tanzania       | 82.0                          | 248.9                      |
| Uganda         | 102.7                         | 130.3                      |
| U.K.           | 1415.7                        | 1120.3                     |
| U.S.           | 8554.9                        | 4994.2                     |

Source: Salamon et al. (2003, pp. 57, 58).

The economic rationale for nonprofit organizations has been considered as the product of demand and supply factors. From a demand perspective, stakeholders may seek supply from an organizational form other than a public or private sector enterprise. This may be due to factors such as ideological motivation, inadequate supply (market failure) or contract failure, “the inability to police producers by ordinary contractual devices” (Hansmann, 1980, p. 835).

Purchasers may consume the service personally, or pay for consumption by someone else, in which case the nonprofit may be a superior provider when service delivery cannot be observed without further cost (Easley and O'Hara, 1983). Although some nonprofit organizations, such as mutual benefit societies, deliver services to specific individuals/groups, their benefits are also consumed by society at large (Andreoni, 1990).

In conditions of market failure (Ben-Ner and Van Hoomissen, 1991)—and the inability of public and private institutions to deliver agreeable solutions to the problems of poverty (Giloith, 1988)—the nonprofit sector operates in a residual role. Market failure could be due to insufficient resources to satisfy heterogeneous demand, economic infeasibility due to inadequate anticipated returns and/or inability to accurately identify demand due to information asymmetry.

Nonprofit organizations have the potential to draw on market and nonmarket resources, such as volunteers, to meet disparate demand. Their lack of profit motive also enables them to draw on resources to subsidize their activities in uneconomic markets and their close engagement with stakeholders better equips them with knowledge and understanding of individual client needs. In this way, nonprofit organizations exhibit differential capability from alternative suppliers through: their greater understanding of, and flexibility for satisfying, individual and disparate client and community needs, particularly those of marginalized groups (Kramer, 1987); their ability to generate indirect benefits over and above their direct outputs; and their capacity to mobilize volunteer labor and donations. The nonprofit sector can thus be more flexible than both government agencies constrained by legislative mandates and for-profit organizations constrained by the economic pressures of the market (Rose-Ackerman, 1996).

In addition to market failure, the existence of nonprofit organizations has been explained by conditions of contract failure. Contract failure (Hansmann, 1980) arises from information asymmetry due to the separation of purchasers and consumers, leaving the seller in a better position to be informed about dimensions of service delivery than the purchaser (Weisbrod, 1989). A consumer in Oxfam, for example, is separated from the ultimate beneficiary of his purchase, and has to rely on assurances from Oxfam that the intended benefit he has contributed to will be delivered to the ultimate beneficiary. In a nonprofit organization, however, the NDC weakens the incentive not to pass on the contributions to the ultimate beneficiary, and therefore prevents the seller from gaining pecuniary advantage from information asymmetry.

Nonprofit organizations have also been justified as more efficient providers of nonrival, nonexcludable public goods (Hansmann, 1980; Ben-Ner and Van Hoomissen, 1991). Public goods, such as community safety and security, cost no more to provide to one consumer than to many. They are also

indivisible: when supplying individual consumers, it is not possible to prevent others from consuming the benefit too. For the market to operate efficiently, each individual should contribute the sum equal to the value he places upon it. Individuals, however, have little incentive to contribute since their individual contribution is likely to be small in proportion to the amount of the service that is supplied, and they can enjoy the benefit of the service that is financed by the donations of others. The problem of “free riding” (Gassler, 1990) arises when little of the good is produced yet collective demand is high. While this is generally solved by state provision of public goods, political ideology and resource constraints may limit government supply of some public goods. Nonprofit organizations then enter the market to supply a public good on a nongovernmental basis, and their nonprofit status provides assurances to the public that donations will be put to use to create social value.

Delivering services, such as community safety and security, through nonprofit social enterprises generates additional gains. These include involving local people in contributing more to revitalizing communities and empowering citizens; encouraging public trust and confidence; helping the sector become more effective and efficient, maximizing social and economic potential; and enabling the sector to become a more active partner with the government in shaping policy and delivery (Gov, 2002).

The Goodwin Center, located on an inner-city estate in the North East of England, is frequently cited as an example of how community activism can turn an estate around. The estate was once notorious for crime, drugs, and prostitution—but now has the Center that houses conference rooms, a café, an IT suite and the usual community facilities, such as meeting rooms. The estate introduced the first community-led warden initiative in the U.K., installing a revolutionary, wireless CCTV system in which digital pictures are beamed back to a control center and can be instantly emailed to the police via the Internet. This scheme is being marketed to other local authorities and any surplus revenue generated will cross-subsidize further regeneration work on the estate. Local stakeholders are actively involved, and all trustees are from the estate and are elected at the annual general meeting (DTA, 2003).

The demand for nonprofit organizations may also be ideological (Rose-Ackerman, 1996): some consumers actively seek to purchase from nonprofit organizations as a result of personal values and beliefs. The preference may be idiosyncratic, or grounded in perceived shortcomings of suppliers from the public or private sectors, such as excessive bureaucracy, inflexibility and inefficiency. Other reasons could include: awareness of the profit motive operating as an incentive to deceive (Ben-Ner and Hoomissen, 1991); incomplete contracts (Glaeser and Shleifer, 1998); information asymmetry (Weisbrod, 1989; Francois, 2003); and difficulty in monitoring the marginal impact of contribution (Weisbrod, 1988; Ben-Ner and Van Hoomissen, 1991; Gui, 1991).

The NDC operates as a protection device for consumers, donors and employees. It provides assurances and builds organizational legitimacy and consumer confidence (Frumkin, 2002). Its nonprofit status ensures that the incentive to generate a surplus (perhaps by reducing quality) is weakened (Glaeser and Shleifer, 2001). In addition, it helps to address fear of exploitation (Rose-Ackerman, 1996) by providing assurances that contributions, and any excess revenue, will be used to further the mission of the organization—they will not be converted into profit (Francois, 2003) or appropriated by the organization's residual claimants (Bilodeau and Slivinski, 1998).

In 1998, the Divine fair trade milk chocolate company was established by a coalition of partners comprising of the Co-op (U.K.), Twin Trading, the Body Shop, Christian Aid and Comic Relief. The organization pays a guaranteed fair trade price to cocoa farmers in Ghana and manufactures a branded chocolate bar which is sold in major retail outlets in the U.K., such as Tesco, Waitrose, the Co-op and Oxfam. The company also sells own label chocolate products to Starbucks and the Co-op, and funds a range of social projects via the farmers' trust. By 2003, the turnover of the company was £2.1 million and its products were sold in 15,000 retail outlets. ([www.divinechocolate.com](http://www.divinechocolate.com))

In markets where for-profits and nonprofits compete, there is some evidence of systematic differences in cost and price behavior between providers. Weisbrod (1989) reported that when hospital payments were made by insurers on the basis of costs incurred by the hospital, for-profits had higher expenses than nonprofits; and when payments were fixed on a *per diem* rate, for-profit hospitals had lower costs per day than nonprofit ones. Glaeser and Shleifer (2001) refer to the higher use of sedatives in for-profit care homes (which make care of the patient easier), and in the health care industry, nonprofit providers have been found to deliver more and better services and produce higher levels of patient satisfaction than for-profit providers (Weisbrod, 2004). Thus, in some markets, nonprofits have been found to produce better outcomes than for-profits (Easley and O'Hara, 1988).

The presence of nonprofit organizations can also be considered from a supply perspective. In the private sector, the motivation for the entrepreneur to create an organization is to make a profit that they can appropriate (Ben-Ner and Van Hoomissen, 1991). Even though the NDC weakens the force of this motivation for nonprofit organizations, there is a consistent flow of new nonprofit organizations and many entrepreneurs create such organizations (Glaeser and Shleifer, 1998).

A nonprofit organization will be created when "the expected flow of net benefits to the entrepreneur is greater than the benefits they can derive from other sources" (Ben-Ner and Van Hoomissen, 1991, p. 532). Its creation will therefore be dependent on establishment costs and the nature and flow of benefits derived from its creation, which must be different in type from

those that flow from the creation of a for-profit firm (Bilodeau and Slivinski, 1998). The lack of financial incentives has led researchers to consider “the power of ideas and emotions in motivating behavior” (Rose-Ackerman, 1996, p. 701), in particular the desire for stakeholders to seek control and advance their interests (Ben-Ner and Van Hoomissen, 1991), altruism (Andreoni, 1990; Rose-Ackerman, 1996, 1997) or psychic income (Gui, 1991).

Ben-Ner and Van Hoomissen (1991) explain the supply of nonprofit organizations from a development perspective and propose that their creators are synonymous with consumers who innovated to fill a market gap (Parker, 2004). Nonprofit status may also overcome information asymmetries between suppliers and customers, such as when for-profit status might “hinder the organizations’ effectiveness by undermining trust between the organization and its stakeholders” (Oberfield and Dees, 1994, p. 566). The creation of a nonprofit organization may also be related to altruistic motives (Drucker, 1990; Rose-Ackerman, 1996) in which the entrepreneur derives benefit from selling services at below-market price to gain enhanced status, or from generating benefits for recipients to produce a “warm glow” from his actions (Andreoni, 1990). Theoretically, the pure altruist derives no personal benefit from the act of giving and cares only about the benefits created for others. Although different forms of altruism exist (behavioral, motivational and environmental) (Wolfe, 1998), proof of the relationship between altruism and the creation of a nonprofit organization remains illusive.

C.O.P.E. (Community Opportunities for Participation Enterprise) was established by Frank Millsop in 1997 in the Shetland Islands, north of Scotland. By 2003, it provided employment opportunities for more than 40 staff, trainees, and volunteers. Frank’s motivation to set up the organization lay in his belief that it was possible to create a business in which people with learning disabilities could be involved. The business includes several different activities, such as catering, retailing, coffee roasting and whole foods packaging. Turnover in 2003 was £230,000, and a separate trading subsidiary, Shetland Soap Company, achieved sales of £100,000. C.O.P.E. has also established a charitable trust to support community activities with funds from the oil industry, which has a presence on the Islands ([www.cope.shetland.co.uk](http://www.cope.shetland.co.uk)).

Another possible reason for adopting the nonprofit format may be in order to obtain resources and advantages not available to for-profit organizations (Bilodeau and Slivinski, 1998). Nonprofit status may enable an organization to acquire resources at a preferential rate, to secure favorable tax incentives, engage volunteer labor and receive donated resources. Although the evidence to date suggests that tax incentives do not explain nonprofit status (Hansmann, 1980; Chang and Tuckman, 1990; Glaeser and Shleifer, 2001), when the provision of a public good is financed by voluntary contributions, the ability to leverage resources has been found to be enhanced by nonprofit status (Bilodeau

and Slivinski, 1998). In these circumstances, it may be in the entrepreneur's interest to establish a nonprofit enterprise.

In summary, the establishment of a nonprofit organization is the product of confluence of demand and supply factors (James, 1990; Ben-Ner and Hoomissen, 1991), the force of which is likely to vary by country, by sector and by individual motivation. Academic, practitioner and media interest in nonprofit organizations has re-emerged in the twenty first century, partly influenced by their potential to provide new social services, generate jobs and promote social cohesion (Defourny, 2001). Many nonprofit organizations are, however, encouraged to adopt a more entrepreneurial approach to doing business due to the changing role of the state from a deliverer to a facilitator via service delivery contracting (Goerke, 2003); increased competition for grants and donations and the fall in funds available for distribution to nonprofit organizations by philanthropic organizations (Jas et al., 2002, p. 11); increased consumer activism and the fair trade movement; and the availability of new forms of social finance. The chapter now moves on to consider nonprofit social entrepreneurship.

## 5. NONPROFIT SOCIAL ENTREPRENEURSHIP

Nonprofit organizations have been categorized as either donative or commercial (Hansmann, 1980) according to their source(s) of income. Donative nonprofits obtain their funds from donations and philanthropy, whereas commercial nonprofits generate at least some of their revenue from trading. If they are in competition with other nonprofit and/or for-profit organizations for resources and customers (Hansmann, 1980; Steinberg, 1993), then their tax and fiscal benefits (Glaeser and Shleifer, 2001) and close stakeholder relationships have the potential to be exploited to generate competitive advantage.

By the end of the twentieth century, the trend was for nonprofit organizations to become more entrepreneurial (Young, 1997), to adopt more innovative strategies and techniques, and to be engaged in generating revenue to create social value, which lessened dependence on philanthropy and donations (Dees, 1998; Weisbrod, 1998b; Reis and Clohesy, 2001). Nonprofit organizations—characterized by their entrepreneurial outlook, autonomy and the pursuit of financial sustainability through the production and sale of innovative goods and services to satisfy the demands of a community that are not met by the State or by the market—have collectively been referred to as social enterprises (OECD, 1999) or social purpose enterprises (Wallace, 1999).

To illustrate, Emmaus is a worldwide secular movement that began in 1945 in France and now operates in 45 countries. The movement is dedicated to the principle of “helping homeless people help themselves.” It is structured around the creation of living and working communities that offer homeless



people a place in a caring, working place where they can feel safe and secure while learning the skills necessary to live an independent life. Each community has its own moneymaking business(es) enabling it to be financially independent. Most communities earn money from the collection and resale of donated goods, refurbished and repaired furniture and electrical goods and recycling projects. After paying their own costs and expenses, the Emmaus philosophy is sustained by using any surplus to help others in need ([www.emmaus.org.uk](http://www.emmaus.org.uk)).

The commercial transformation of nonprofit organizations has served to professionalize management (Alexander, 2000), focus their performance on the “bottom line” (Frumkin, 2002, p. 152), and encourage them to adopt measures to evaluate their social and economic gains (OECD, 2003). To capture their role in social value creation, performance indicators and the concept of the bottom line have been expanded to include social (double bottom line) and environmental (triple bottom line) outcomes. This trend has attracted the attention of policy makers and practitioners who are interested in the potential contribution of social enterprises to economic, social and/or environmental regeneration and renewal.

Although there is no universally agreed definition of a social enterprise, there appears to be a general consensus that it is a business with primarily social and/or environmental objectives, whose surpluses are principally reinvested for that purpose either in the business and/or a community rather than being driven by the need to maximize profits for shareholders and owners (DTI, 2002). Nonprofit social enterprises can be distinguished from other nonprofits by their strategies, structure and values (Dart, 2004). They can be distinguished from for-profit social enterprises by their retention of any financial surplus.

Social enterprises are engaged in revenue generation from trading (Zeitlow, 2001) with a benchmark of 50% and above (DTI, 2002). Some examples of earned income strategies include fees for service, sale of goods, service delivery contracts, course and tuition fees, consulting, rental income, lease payments (Skloot, 1987; Frumkin, 2002; Zeitlow, 2001) and cause related marketing (Varadarajan and Menon 1988; Weisbrod, 2004). Earned income strategies are increasingly used by charities—in 2002, for example, contracts and trading accounted for approximately 33% of total income for charities in the U.K. (NCVO, 2002). This trend is forecast to continue as nonprofit organizations become more involved in delivering statutory services under contract.

Sunderland Care Home Associates (SCHA) is a major contractor to the City of Sunderland’s social service department. SCHA employs 150 people and provides personal care and domestic services (mobility, washing, dressing, feeding and cleaning) to home-based clients. It also provides disability support services to students at the local university. The philosophy of SCHA is to promote independent living: the services they deliver enable old, frail and

TABLE 14-3 *Civil society sector FTE revenue*

|              | Government<br>(%) | Philanthropy<br>(%) | Fees, dues<br>(%) |
|--------------|-------------------|---------------------|-------------------|
| Mexico       | 8.5               | 6.3                 | 85.2 (1)          |
| Kenya        | 4.8               | 14.2                | 81.0 (2)          |
| Brazil       | 15.5              | 10.7                | 73.8 (3)          |
| Argentina    | 19.5              | 7.5                 | 73.1 (4)          |
| Colombia     | 14.9              | 14.9                | 70.2 (5)          |
| Ireland      | 77.2 (1)          | 7.0                 | 15.8              |
| Belgium      | 76.8 (2)          | 4.7                 | 18.6              |
| Germany      | 64.3 (3)          | 3.4                 | 32.3              |
| Israel       | 63.9 (4)          | 10.2                | 25.8              |
| Netherlands  | 59.0 (5)          | 2.4                 | 38.6              |
| Pakistan     | 6.0               | 42.9 (1)            | 51.1              |
| Romania      | 45.0              | 26.5 (2)            | 28.5              |
| Slovakia     | 21.9              | 23.3 (3)            | 54.9              |
| South Africa | 44.2              | 24.2 (4)            | 31.7              |
| Tanzania     | 27.0              | 20.0 (5)            | 53.1              |

Source: Salamon et al. (2004, p. 299, Table A4).

disabled people to live in their own homes for as long as possible. SCHA delivers more than 3500 hours of services a week, has a turnover of more than £1 million per year and is financially sustainable (SEC, 2003).

The lack of international consensus concerning their definitive characteristics (Sullivan et al., 2003) makes precise, comparable data on social enterprises difficult to locate at national and international levels. Assuming that revenue generation is linked to entrepreneurial activity, an indication of their presence can be derived from the John Hopkins Comparative Research Project. Of the 32 countries where revenue data was available overall, 53% of income was generated from fees, charges for services, related commercial income, and other commercial sources; 35% from government or public sector support (grants, contracts and reimbursement of payments); and 12% from private giving (donations and philanthropy).

This data reveals important country and industry differences in dependence on income source. Table 14.3 summarizes the percentage of non-profit organizations according to income source and shows that reliance on independent revenue is greatest in Mexico (85.2%), Kenya (81.0%), Brazil (73.8%), Argentina (73.1%) and Colombia (71.4%). It is least in Ireland (15.8%), Belgium (18.6%), Israel (25.8%), Romania (28.5%) and South Africa (31.7%) (Salamon et al., 2003). In relation to industry sector, on average, revenue generation is important (accounting for more than 50% of income) for professional, culture, development and housing, foundations, education and other nonprofit organizations.

Country level studies also provide an insight into social enterprise activity. In a study of 105 nonprofit business ventures from 72 organizations in the U.S., 69% reported that their ventures broke even, taking on average 2.5 years to reach that point. The majority generated modest revenues, 27% generated less than \$100,000 and 17% generated more than \$1 million (Community Wealth Ventures, 2003). Borzaga and Defourny (2001) review social enterprise activity in 15 European countries, but do not include comparable, quantitative data on the sector. In the U.K., the Social Enterprise Coalition (SEC, 2003) estimated that there were more than 5000 social enterprises that earned more than 50% of their income from trading and employed nearly 5.5 million people (Gov, 2002, p. 24). More recent data (SBS, 2005) has increased the estimate to approximately 15,000 social enterprises that are registered as Companies Limited by Guarantee or Industrial and Provident Societies, 88% of which generate 50% or more of their income from trading.

The legal format adopted by a nonprofit social enterprise will vary from country to country, influenced by the type of activity, legislative constraints and other country-specific factors (OECD, 1999). In the U.S., for example, the term social enterprise generally applies to a social-purpose venture run by a nonprofit organization to generate a surplus which is used to finance their nonprofit activities (OECD, 1999). Similarly, in the U.K., the range of alternative legal constitutions means that social enterprises can appear on a number of different registers (such as those held by the Charity Commission, Companies House or the Register of Industrial and Provident Societies).

The collation of reliable data is further complicated by the lack of agreement concerning definition, particularly in relation to the extent of revenue generation. Dees (1996, 1998) defined social enterprises on a continuum stretching from those that are entirely commercial and market-driven to those that are entirely charitable, donative and voluntary. Others assert that social enterprises must generate some revenue from trading (Boschee and McClurg, 2003). To capture the variation in extent of revenue generation, the DTI (2002) has recently broadened its definition to include “emerging social enterprises” that generate between 25–50% trading revenue. Although it is difficult to prove at present—and the importance and extent of revenue generation varies between countries and sectors—there appears to be a general consensus that social enterprise activity may be increasing.

The social purpose of a venture may be to provide a private good for individual benefit, such as affirmative employment, or to provide a public good, such as alleviating poverty, pollution and human suffering (Wallace, 1999). The structure of the social enterprise might combine a range of activities, or projects, with the surplus from the more profitable ones being used to cross-subsidize those less profitable (Schiff and Weisbrod, 1991). This may be achieved by the creation of a subsidiary of activity related, or unrelated, to the core mission

of the organization. Country-specific regulations will apply to the taxation of related and unrelated activity.

The Delancey Street concept—established in San Francisco in 1971 by four people and a loan of \$1000—illustrates this approach. The aim of the organization is to enable substance abusers and former convicts to rebuild their lives in a structured educational and living environment. By 2003, the organization owned a large complex with accommodation for 500 disadvantaged people. In common with Emmaus, Delancey Street manages several social purpose, revenue generating activities in which individuals receive training to help them lead an independent life. These include a restaurant, a café, a bookshop, a removal company, automotive services and the retail of Christmas products. Each person stays an average of four years, during which time they gain education and work skills, while learning accountability, responsibility, dignity and integrity. Since its establishment, Delancey Street has built and remodeled 1500 units of low-income housing, trained 800 people in building skills, 10,000 formerly illiterate people have gained high school equivalent degrees and 14,000 people have become tax paying citizens (<http://www.eisenhowerfoundation.org/grassroots/delancey/>).

The nonprofit status of a social enterprise enables the pursuit of activities that would be difficult for a for-profit organization to achieve. An example would be in the case where it is not feasible to charge a fee that covers all the costs of providing a product/service. The goal of financial sustainability when set against the cost of supply and level of demand, however, is likely to make certain market opportunities more attractive than others. This may tempt some social enterprises to concentrate on the most appealing, satisfying and manageable causes, leaving the most difficult work undone (Frumkin, 2002, p. 140). It is also likely to impact on the civil society role of social enterprises in that, by relying on commercial markets, their need to build networks with nonmarket stakeholders is reduced (Eikenberry and Kluver, 2004). Furthermore, social enterprises may be pushed into diverting scarce resources away from their social purpose to compete with other organizations attracted by the market opportunity. For these reasons, generating revenue from trading may not be suitable, or achievable, for all social enterprises. It may be rejected if the strategy is perceived as either compromising social purpose (Weisbrod, 2004), disruptive and/or not in the organization's interest (Skloot, 1987).

## 6. THE SOCIAL ENTREPRENEUR

From a functional perspective, the supply of new enterprises to an economy is determined by the actions of entrepreneurs who organize resources to create new sources of supply in order to meet changes in demand (Badelt,

1997). Early research sought to explain entrepreneurial behavior by analyzing the characteristics and personality of the entrepreneur, for example, by describing him as self-centered, power seeking and independent (Schumpeter, 1934). This approach did not produce conclusive results (Chell et al., 1991) and has for the most part been abandoned (Shaver and Scott, 1991) due to its inability to identify characteristics that distinguish entrepreneurs from nonentrepreneurs.

In a similar pattern, the motivation to create a social enterprise has been linked to personality traits (Leadbeater, 1997; Thompson et al., 2000; Shaw et al., 2001; Johnson, 2003). For the nonprofit entrepreneur, motivations may be based on personal development (the search for personal identity), the belief in a cause/mission (Badelt, 1997; Hibbert et al., 2002) and/or the desire to change society (Bornstein, 2004). In this way, the nonprofit enterprise is a vehicle for the expression of personal values (Young, 1983; Rose-Ackerman, 1996, 1997; James, 1987, 1989; Gassler, 1990; Oberfield and Dees, 1994), virtuous behavior (Mort et al., 2003) and/or spiritual beliefs (Frumkin, 2002, p. 26) in a private vision of doing good. Once again, these claims should be considered cautiously since evidence shows that ideological entrepreneurs may also create for-profit enterprises (Quarter et al., 2003).

In one of the first studies to consider social entrepreneurs, Young (1983) found that they possessed alternative motivational leanings that led them to purposefully choose to create an enterprise in the nonprofit sector. In 1986, he proposed a novel behavioral theory of social entrepreneurship and defined a set of models of nonprofit motivation. They were: artist, professional, believer, searcher, independent, conserver or power-seeker.

Young's theory proposed that entrepreneurs would gravitate to specific sectors depending on the intrinsic nature of the services, the degree of professional control, the level of industry concentration and the social priority of the field. Entrepreneurs would find their place in either the public, private or nonprofit sector in a process driven by: the desire to realize income; the level of hierarchy and bureaucracy acceptable; and orientation toward service (Young, 1986). In retrospect, this theory modeled entrepreneurship narrowly: it is likely that individuals combine several traits and motivations in their activities (Frumkin, 2002). The recent social enterprise spectrum (Dees, 1996) is more accommodating of the complexity of motivational drivers.

Social entrepreneurs have been described as "path breakers with a powerful new idea" who are "totally possessed" (Bornstein, 1998, p. 35) by their vision. They are change agents (Waddock and Post, 1991), pioneers who pay attention to market forces without losing sight of their social mission (Boschee, 1998). Their motivation may take some form of altruism (Piliavin and Charny, 1990)—achieving personal satisfaction from one's own act of charity, the "warm glow" (Andreoni, 1990), or gaining satisfaction from the gratitude and affection of beneficiaries (Rose-Ackerman, 1996).

More broadly, Oberfeld and Dees (1994, p. 566) noted that a common element shared by social entrepreneurs was the desire to “make a contribution to society” in association with a belief that the contribution could not be made “through a traditional business venture.” The studies referred to here, however, are case study based, and the reliability and validity of findings would be improved if control groups were used for future studies.

Tim Smit is a social entrepreneur, best known as the co-founder and driving force behind the Eden Project in Cornwall. This Project converted a former china clay pit near St Austell into the world’s largest greenhouse, consisting of a series of functioning eco-systems (biomes). Its aim was to create wealth for the benefit of the wider community: the venue employs 390 staff, mostly from Cornwall and is the third most popular tourist attraction in the U.K.

In addition to this, the Eden Project has contributed to local development by sourcing locally wherever possible. It has boosted the performance of local businesses (92% of visitors stay in holiday accommodation) and an economic study valued its annualized impact at £150 million (*www.edenproject.com*).

Weisbrod (1988) proposes that the advantages of nonprofits (low barriers to entry (Frumkin, 2002:130), access to tax and fiscal benefits and resources not available to for-profit enterprises, and trade from ideology-led customers) may result in some for-profits disguising themselves as nonprofits and masking their private profit-seeking activities. Furthermore, any weak enforcement of the NDC will enable profits to be diverted away from the social purpose of the organization (Bilodeau and Slivinski, 1998) and into internal improvements, such as enhanced resources from increasing endowment, and employee benefits, such as “lower effort levels, free meals, shorter hours, longer holidays, better offices, more generous benefits” (Glaeser and Shleifer, 1998). This shows that the theme of motivation to create a nonprofit social enterprise needs to be considered more critically than has been to date in social entrepreneurship literature.

At present, internationally comparable, comprehensive data reporting the characteristics and motivations of social entrepreneurs is not available. In a unique study, Harding and Cowling (2004) investigated social entrepreneurship in the U.K. and found that 6.6% of the U.K. population was engaged in some form of activity that had a social or community goal, either as a start-up or as an owner-manager of that venture. This rate is slightly higher than overall levels of for-profit entrepreneurial activity, which was at 6.4% in 2003.

Harding and Cowling found social entrepreneurs to be older, more qualified and employed on higher incomes than mainstream (for-profit) entrepreneurs. They also found high levels of social entrepreneurial activity among disadvantaged groups (low incomes/unemployed/women/ethnic minor-

ity groups). Social entrepreneurs were more positive and less likely to know an entrepreneur. They were also less likely to see good opportunities, to think they had the skills to start a business or to let lack of finance prevent them from starting up their social enterprise. From a business perspective, they created more jobs, were increasingly likely to pay salaries than rely on volunteers and generated higher median turnovers than for-profit entrepreneurs. The survey found that sales revenue accounted for more than 50% of revenue for 58.9% of start-up social enterprises and 47.2% of owner-managed social enterprises. The results have created an important benchmark and comparable, international data would be helpful to extend the research.

The process of creating a new, for-profit venture combines innovation, organization creation and profit-seeking behavior (Hornaday, 1992). Social enterprise creation has been found to differ from for-profit creation in terms of the pursuit of social value, the reliance on a mixture of employees and volunteers, multiple sources of income of which earned income is just one component, evaluation of performance in terms of social value, tax benefits and the NDC (Skloot, 1987). The final part of this chapter considers the process of social enterprise creation: opportunity recognition, innovation, resource acquisition, opportunity exploitation, venture creation and outcomes.

## 7. VENTURE CREATION

### *7.1. Opportunity Recognition*

A new venture is created to exploit an opportunity, the potential value of which is perceived and acted upon differently by entrepreneurs and potential entrepreneurs (Casson, 1999). Kirzner (1997) refers to entrepreneurial alertness: the propensity of the entrepreneur to notice, without searching, the opportunities that have been overlooked by others. Drucker (1985) identifies three categories of opportunity: inefficiencies in markets, changes in trends and inventions/discoveries. Market inefficiencies due to market and contract failure—and changing economic, social and political trends—have created opportunities for social enterprises, as have advances in technology. Catford has summarized the importance of entrepreneurial alertness and opportunity recognition in social entrepreneurship. He describes social entrepreneurs as individuals who “see opportunities where others only see empty buildings, unemployable people and unvalued resources” (1998, p. 96).

The opportunities for social value creation embrace a wide range of potential organizational objectives, several of which may be pursued simultaneously. These include reducing social, economic and financial exclusion; reintegrating the long-term unemployed back into the workforce; generating

innovative and dynamic solutions to revitalize deprived areas; and acting as a vehicle for social cohesion and as a place of socialization.

The increasing demand for social services (Salamon, 1993) has created many new opportunities, although it is perhaps unrealistic to expect the sector to be able to satisfy the market as its capacity tends to be limited by resource constraints. Strategically, serious consideration will need to be given to which opportunities should be pursued and which would be better handled by other organizations, especially since some opportunities may be attractive to for-profits and nonprofits alike (Sloan, 1998; Tuckman, 1998; Ryan, 1999). Competition between for-profits and nonprofits has already been noted in health care (Clarke and Estes, 1992), care of the elderly (Mort et al., 2003), childcare (Kagan, 1991), and welfare to work (Frumkin and Andre-Clark, 2000).

## 7.2. *Innovation*

The opportunity to create social value can be considered from the perspective of innovation. In the for-profit sector, innovation is a fundamental element of entrepreneurship (Schumpeter, 1934; Drucker, 1985). Schumpeter (1934) identified five categories of innovation: the introduction of a new economic good or service; the introduction of a new method of production; the opening of a new market; the conquest of a new source of raw materials; and the reorganization of an industry, such as the creation or dismantling of a monopoly. Each of these innovation opportunities offers a potential market in which social enterprises could create social value—for example, Pearce (2003) identified the following market opportunities in disadvantaged communities: local development and regeneration through managed workspace, business incubators, enterprise training and business advice and support; service contract delivery on behalf of the government; filling market gaps; and market-driven businesses in direct competition with the public and private sectors.

Although innovation is generally underexplored by nonprofit organizations (Zimmerman, 1999), nonprofits have been innovative in delivering services to markets (Kramer, 1987) and developing new forms of supply (Moss Kanter, 1999). More recent developments in innovation theory have proposed that the future of competition will focus on the co-creation of value in relations between the organization and its customers (Pralahad and Ramaswamy, 2004). In this respect, the close relationship between social enterprises and their stakeholders, in some cases being both producer and consumer, presents them with a potential competitive advantage over other organizations (Leadbeater, 2004).

An example of this is Green-works, an office furniture refurbishment and recycling venture. It brings together corporate organizations (such as HSBC, Unilever, Marks and Spencer) wanting to dispose of office furniture and



community groups that need it. The business has a turnover of approximately £1 million year, and has recycled more than 3000 tons of furniture. Over 600 community groups, schools and charities have acquired furniture at low prices. Green-works has created employment opportunities for people who are homeless, long-term unemployed or from disadvantaged backgrounds ([www.green-works.co.uk](http://www.green-works.co.uk)).

### 7.3. Resource Acquisition

The process of creating a new venture involves the gathering and combining of resources. For social enterprises, five major categories of resources—physical, human, financial, technological and other (such as reputation and legitimacy)—may be acquired through donation, purchase and/or barter (Reisman, 1991). A brief overview of these categories is useful to explain the distinctiveness of resource base expansion in the nonprofit social enterprise.

Physical resources (property, land, vehicles, and equipment) are required to establish a base for the organization and create an infrastructure for social value delivery. The nonprofit status of social enterprises may give access to donated assets or the ability to purchase assets at below market price. The strategy of asset base development has been promoted as a means of achieving financial sustainability through community regeneration (Hart, 2001). Investment appraisal of potential assets is essential, however, since some may be acquired cheaply but require massive expenditure to make them operational.

Coin Street Community Builders (CCSB) is a nonprofit development trust located on the South Bank, Central London. In 1987, CCSB took ownership from the former Greater London Authority of a 13-acre site that borders the River Thames. The most famous landmark of the site is the OXO Tower, which stands on the top of the main building. The site was acquired at below market value and a restrictive covenant attached to its development.

Since acquiring it, CCSB has developed housing for local people and key workers, as well as commercial and recreation facilities. The site was developed with a combination of grants and loans, which were used to fund asset base development. It now includes two exclusive restaurants that are leased to a major retailer, craft shops, parks and four cooperative social housing projects that will eventually house 1300 people (<http://www.coinstreet.org/indexIE.html>).

The physical effort to manage an organization, and create and deliver the social value proposition, is supplied by human resources—stakeholders, trustees, employees and volunteers. Since nonprofit organizations cannot compel action and rely on noncoercive participation or consumption of their outputs (Frumkin, 2002), it has been proposed that human resources are drawn to them by their mission or purpose (Oberfield and Dees, 1994; Francois, 2003).

Preston (1988) found that employees in the nonprofit sector valued job quality more and wages less than for-profit employees. In some cases they may accept lower levels of pay in return for achieving their altruistic goals (Preston, 1989). Perceived wage differentials between for-profit and nonprofit enterprises (Knapp, 1989; Preston, 1993) has, however, led to difficulties in attracting and retaining sufficiently qualified and skilled employees (Zeitlow, 2001). This deficit is likely to become more pronounced as nonprofit social enterprises increasingly seek out employees with skills in revenue generation.

In addition to employees, nonprofit social enterprises may draw on free volunteer labor. The knowledge and networks that they bring may reduce transaction costs. The involvement of volunteers may also enhance legitimacy through the relationships between the enterprise, the community served, and society more generally.

Increasing the market focus of a social enterprise, however, may produce a negative impact by reducing volunteer involvement. This was recorded in a study of voluntary sport organizations by Enjolras (2002): it was found that nonprofessionally active members decreased their amount of voluntary work as commercial income increased.

Although nonprofit social enterprises lack access to equity markets, they have access to five principal sources of income: private donations/philanthropy, government grants, trading revenue (either from contracts with public or private sector organizations or from selling goods and services), debt finance and investment income (either from dividends, interest or rents from investment property). Most private giving goes to nonprofit organizations (Rose-Ackerman, 1996) and it would be difficult to imagine private donations to a for-profit firm (Glaeser and Shleifer, 1998). The rationale for private donations and philanthropy is personal to the donor and may reflect their whims and personal goals (Oberfield and Dees, 1994). Any increase in emphasis on revenue generation is likely to reduce dependence on donations/philanthropy which is still an important source of finance for many organizations (Etherington, 2004).

The Big Issue was launched as a nonprofit enterprise in London in 1991. Although re-constituted as a for-profit venture, its core mission remains the same: to enable homeless individuals to work and earn income and to campaign on social exclusion issues. Homeless people are provided the opportunity to earn income from selling magazines on the basis of quality, not as a means of soliciting a charitable donation. The organization campaigns for social welfare and raises funds for the Big Issue Foundation (Hibbert et al., 2002).

The benefits of revenue generation are that it confers greater autonomy (Alexander, 2000), independence and control over a social enterprise's resources. On the other hand, the amount of potential revenue to be generated from trading will be determined by supply and demand, and those social enterprises that serve thin, nonmainstream markets may find it difficult to

achieve sufficient income to remain solvent. Additionally, retaining sufficient funds from trading to finance major capital expenditure may not be feasible in the short term, and this will impact on future capital investment and growth.

Theoretically, nonprofit organizations and social enterprises have access to the same sources of finance as for-profit enterprises. Their uptake of loans, however, has been low (Bank of England, 2003). This can be explained by high transaction costs (bureaucracy), information asymmetry, having low priority with lenders, lack of critical mass and few proven finance vehicles. These factors have led to the creation of community and social finance institutions specifically for lending to social and social-purpose organizations (OECD, 1999; Kingston and Bolton, 2004).

In the U.S., the Community Reinvestment Act (1977) was introduced to support the involvement of financial institutions in responding to the needs of deprived communities (OECD, 1999). There have also been developments in community development venture capital (CDVC) (Jegen, 1998) and venture philanthropy (VP). VP combines venture capital principles with the practice of philanthropy (Letts et al., 1997), considering grants to organizations as an investment, supplemented by sustained advice and support.

In the U.K., The Community Development Banking and Financial Institutions Act (1994) established guidelines for the creation of Community Development Financial Institutions (CDFIs) to finance regeneration initiatives. The Social Investment Task Force (<http://www.enterprising-communities.org.uk/>) recommended the introduction of a community investment tax credit to stimulate investment in social purpose enterprises. In 2003, The Charity Bank was launched to provide patient loan finance to social and community purpose enterprises. The share of income generated by general charities increased between 1995–2001 (Jas et al., 2002, p. 31), and it is likely that the demand for these new financial products will also increase (Jegen, 1998).

The relationship between different sources of revenue is worth noting. Evidence of increasing sales revenue, for instance, may lead to crowding out in that revenue from other sources, such as donations, falls as sales revenue increases. Segal and Weisbrod (1998) noted this, finding that donations and sales revenues were substitutes for each other in nonprofit firms. In addition, an increase in commercial income has been construed as conferring unfair competitive advantage to the nonprofit social enterprise (Schiff and Weisbrod, 1991).

The YMCA has been criticized in this regard. Established in 1851, it provides health and fitness facilities across the world and generating revenue of \$4.1 billion in 2001. Its charitable status has meant that it can undercut private health club prices and its increasing presence in more affluent areas has led to allegations that YMCAs have abandoned their mission to serve low-income neighborhoods. In response, the YMCA asserts that its mission is to serve

all income neighborhoods and that the facilities in high-income areas cross-subsidize low-income families (Weisbrod, 2004).

Social enterprises also acquire the resources of information technology (IT), both for infrastructure development and service delivery (Schneider, 2003; Elliott et al., 1998). IT resources are employed to create the organization, to facilitate the construction of virtual social enterprises, and to establish online communities through which to foster economic and social inclusion. IT skills training forms a major part of social enterprise activity, especially delivering IT training to enable the unemployed to join the workforce. In addition, IT is employed to contribute to individual and societal capacity building, for example, enabling hard-to-reach individuals and communities to access information and employment opportunities online.

The reputation related benefits and legitimacy of nonprofit social enterprises might also be construed as a resource: they communicate a collection of socially-constructed values. These values include offering a higher quality service, being more trustworthy, operating with greater integrity and being more client-focused. Although unverifiable and noncontractible (Glaeser and Shliefer, 1998), they may be used to generate income from trading contracts and/or from cause-related marketing initiatives (File and Prince, 1995).

The nonprofit status of the social enterprise thus confers resource advantages as well as disadvantages. The attractiveness of its social purpose and tax incentives reduce outgoings, and revenue generation public donations of funds, time and assets increase resources. Resources may be acquired directly from the market place either at, or below, market price, or indirectly from networks. If social enterprises grow and adopt the practices of for-profit ventures at the expense of their social goals, however, their resource advantages may be withdrawn or disappear. It is therefore critical that opportunity exploitation and venture creation maintain the purpose and philosophy of the social enterprise.

#### *7.4. Opportunity Exploitation and Venture Creation*

Opportunity exploitation is frequently presented as a sequential process in which the resources that have been gathered are combined to establish a new venture. Typical stages would be: the enterprise is legally constituted, an organizational structure is created and then a strategy is developed and implemented. In social enterprises, this process is managed while balancing the economic imperative to be financially sustainable with the enterprise's distinctive values of cooperation, mutuality, inclusiveness and stakeholder accountability. Several social enterprise business development models have been created to explain the venture creation process (Dees et al., 2001; Dees et al., 2002; Emerson and Twersky, 1996; Weisbrod, 1998a; Brinckerhoff, 2000). These tend to be

descriptive, dividing the process into discrete stages and identifying critical incidents and common problems.

After the venture has been created, decisions concerning strategy, marketing, governance and performance evaluation need to be made and implemented. For social enterprises, an important strategy for achieving sustainability involves entering into partnerships with other nonprofit organizations, public sector institutions (Goerke, 2003; Shaw, 2003) and for-profit organizations (Andreason, 1996; Austin, 2000; Skloot, 2000). These partnerships bring benefits to both parties in terms of leveraging tangible and/or intangible resources that can be combined to generate collaborative advantage (Huxham, 1996). Future growth options are likely to involve consideration of expansion away from serving local, niche markets to providing mainstream goods and services to regional, national and/or international markets.

## 8. ENTREPRENEURIAL OUTCOMES

In the for-profit sector, businesses report their performance to disclose the financial health and the effectiveness of the organization (Dees et al., 2002). Performance can be assessed by a range of measures, goals and indicators such as standardized techniques of financial accounting and auditing, and uniform measures of profit, market share, share value and shareholder return. Although it has been argued that entrepreneurial outputs include social contribution as well as economic performance (Venkataraman, 1997), performance assessment has tended to focus on documented, economic measures. Relatively little attention has been given to developments in the reporting of social and environmental impacts in for-profit firms ([www.globalreporting.org](http://www.globalreporting.org)).

In the nonprofit sector, these measures are either inappropriate or inadequate since their success is related to the creation of social value, which is particularly resistant to enumeration (Moss Kanter and Summers, 1987; Oberfield and Dees, 1994; Speckbacher, 2003). Since nonprofit social enterprises aim to achieve financial sustainability through the creation of social value, they will generate at least two sets of outcomes on which to evaluate their performance: economic value and social and/or environmental value. Although the practice of translating social outcomes into measurable indicators is not without its critics (Campbell, 2002), the process offers a pragmatic method for reporting social enterprise performance (Morley et al., 2001; Quarter and Richmond, 2001).

### 8.1. Business Outcomes

In theory, the business performance of social enterprises can be assessed by the same measures as those used by for-profit organizations (income,

turnover and employment, for instance). In practice, however, this would not acknowledge income and expenditure specific to social enterprises and ignores less direct outputs and outcomes. Actual expenditure categories are likely to include staff costs; goods and services (rent, travel, subsistence, communications, professional fees, financial services, and miscellaneous, such as stationery); grants and donations; fundraising and publicity; and interest payments and depreciation of capital assets.

Ratio analysis that uses the measures from the for-profit sector may be of value if it is acknowledged that outcomes and performance standards are likely to differ between for-profit and nonprofit organizations. Productivity may be lower as a result of affirmative employment practices, gearing may be lower due to low borrowing and asset constraints, and employment data may be difficult to acquire due to short-term contracts and project-based employment. For these reasons, it is likely that new analytical tools and techniques for social enterprise business performance evaluation are needed by the sector.

## 8.2. *Social Outcomes*

The limitations of financial accounting for nonprofit organizations have created a growing industry in social accounting (Richmond et al., 2003), social auditing, social impact assessment and social return on investment appraisal. Social performance, social value and social impact are ambiguous concepts, which are further complicated by the intangible nature of social outputs, the length of time that they take to emerge, and the difficulty in proving causality between social enterprise outputs and the social value created. In addition, performance measurement is undertaken at organizational level and some measures are likely to be idiosyncratic and developed through an iterative process (Sawhill and Williamson, 2001). The costs of measuring social and/or environmental outcomes will be incurred at organizational level, further increasing the pressure on already tight resource constraints.

The performance measures implemented at organizational level will depend on the goals of the business. Success will be measured in terms of the creation of social value (Oberfield and Dees, 1994) and might be quantified by enumerating savings to public expenditure (Dees et al., 2002). If, however, the social purpose of the venture aims to further the common good (Pearce, 1994; Dees, 1994) through societal gains—such as better access to health care (Herzlinger and Krasker, 1987), economic development, social stability and political cohesion (Wallace, 1999); or improved social well being, social capital, social cohesion and increased relational assets (Borzaga and Defourny, 2001)—then key indicators are likely to be almost impossible to define.

The challenge of creating a method that is capable of enumerating and quantifying social outputs has led to a demand for new measures that ac-

knowledge multiple stakeholders (Moss Kanter and Summers, 1987), multiple outcomes, social impact and added social value (REDF, 2001). In principal, social accounting gathers qualitative data on a range of indicators designed to report on the organization's performance in terms of meeting stakeholder expectations (Quarter et al., 2003). Several social accounting methods have been created and each has its own auditing procedure (Pearce et al., 2000). Social accounting has been used to measure subjective concepts and developments in the techniques have addressed social capital, social justice, social inclusion, participation, regeneration, modernization, civil society, citizenship, employability and social well being. Measures tend, however, to be proxy indicators.

Techniques for assessing multiple outcomes also include the Balanced Score Card (Kaplan and Norton, 1996; Kaplan, 2001) and social return on investment (SROI) (Emerson and Wachowicz, 1999; Emerson, 2001; New Economics Foundation, 2004). These techniques aim to record the value of social outcomes by translating the social objectives of the organization into financial measures. The advantages of standardized techniques lie in their transparency and their ability to generate comparable data useful for benchmarking purposes.

The Rubicon Bakery is a wholesale bakery that provides employment and training opportunities in the food service industry for disadvantaged residents. The Bakery was developed in 1994 and by 2000 achieved sales of \$1.14 million through 100 premium grocery stores. It employs 18 individuals, 80% of whom have come from its own training program. Its success could be seen in an SROI report (REDF, 2000), which calculated that each target employee saved \$16,807 in public expenditure and generated \$2911 in new taxes.

The value of social enterprises at local, regional and national levels, and their economic multiplier effects, however, has yet to be conclusively proven. A recent example of a full cost-benefit analysis of one Emmaus community in the U.K. valued its contribution at £26,652 per companion and at least £613,000 for the wider community. The method gathered quantitative data for actual business costs, imputed business costs and accommodation and living costs. It involved evaluating the benefits, costs and outputs of the community. Values for savings included housing and income benefits, asylum seeker support, drug treatment, legal/justice, healthcare, recycling, skills training and death costs ([www.emmaus.org.uk](http://www.emmaus.org.uk)).

## 9. CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

By the end of the twentieth century, the institutional boundaries between the private, public and nonprofit sectors had blurred, with many nonprofit organizations adopting the entrepreneurial strategies usually associated with for-profit enterprises. Despite this trend, however, many of these organizations

still receive substantial resources from philanthropic sources and via public sector service delivery contracts. The nonprofit social enterprise is located in a market place that lies between the private, public and nonprofit sectors, pursuing hybrid strategies that combine social purpose in delivering either public or private goods, for individual and/or societal benefit, with financially sustainable business ventures. To date, their real contribution to social and economic value creation has been difficult to enumerate; the managerial and operational challenges they face leads to the four key areas of future research outlined below.

The first research theme concerns outcomes and performance. There is an urgent need to establish legitimate and standardized performance measures that will accurately and cost-effectively measure the creation of social and/or environmental value by social enterprises. Many claims have been made about the potential of social enterprises to achieve economic, social, and/or environmental outcomes. Currently, the standardized, quantitative measures (example, e.g., employment, income and turnover) provide an important indication of enterprise performance. Such measures exclude intangible, subjective outcomes that might be equally, if not more, important. These outcomes are difficult to evaluate and ultimately may be impossible to measure quantitatively. There are several researchers working in this domain and it is likely that a set of composite, proxy indicators will eventually be created, enabling social enterprises to prove their success in achieving social and economic gains to stakeholders and policy makers. This will undoubtedly be useful and may enable additional resources to be leveraged by the sector.

The second research theme involves examining the management of the relationship between the simultaneous pursuit of economic, social and/or environmental value. In some environments and markets, the social purpose of the enterprise might be achievable only at the expense of economic outcomes, or might have to be sacrificed in order to achieve financial sustainability. The potential of some markets, in terms of revenue generation, is likely to be greater than others, rendering those enterprises that operate in resource-rich markets more capable of achieving financial sustainability. Other resource-poor markets may remain neglected and suffer from market failure. Market failure, however, is one of the explanations for the emergence of the nonprofit sector and a gap that social enterprises are attempting to fill. It is therefore conceivable that social enterprises, instead of addressing market failure, may simply be engaged in a process of stratifying the nonprofit sector by exploiting the most attractive opportunities, leaving a gap for other, perhaps donative, nonprofit organizations to serve. Research that identified models of successful management of the relationships between different objectives would be valuable for the sector.

The third suggested research theme involves investigating the impact of commercial strategies and revenue generation on the shared values of the



nonprofit social enterprise. While shared values of mutuality, cooperation and stakeholder consultation have traditionally been associated with nonprofit sector organizations, the pursuit of financial sustainability might compromise these values or even sacrifice them in preference for entrepreneurial values associated with rapid decision making and risk taking. In learning the skills and competencies and adopting the practices of for-profit organizations, social enterprises may be tempted to simply become, either individually or in partnership with, for-profit enterprises. This would result in abandoning their role in delivering public goods, especially in relation to fostering social capital and promoting civil society. It is also possible that social enterprises may not maintain their role in countering contract failure, and may no longer be attractive to consumers who purchase from them for ideological reasons. Research could examine this theme through comparisons between nonprofit and for-profit social enterprises with a view to identifying strategies that will enable social enterprises to maintain their values even as they become more entrepreneurial.

The final research theme is associated with organizational growth and development. Many social enterprises are small, created to serve local demand and the extent of local demand may constrain future growth of the organization. When social enterprises have been created in deprived, resource and asset-poor communities, demand is likely to have been limited to the local market. In order to grow, serve a larger market and benefit from economies of scale, however, social enterprises need to evaluate the potential of nonlocal demand and the attractiveness of mainstream markets. They may have to abandon their original purpose and market—but in so doing may recreate the market failure they initially set out to address. Research revealing the conditions under which strategies and processes could achieve financial sustainability in thin markets would be of great benefit to the sector, and may also be transferable to for-profit enterprises.

In conclusion, social entrepreneurship is a multidimensional concept that brings together entrepreneurial behavior and opportunities to create social value. Nonprofit social enterprises aim to balance economic, social and/or environmental aims. Their long-term sustainability rests on delivering multiple outcomes. The advantages of earned income strategies can be to free the social enterprise from the constraints and obligations of philanthropy, while increasing the enterprise's flexibility and adaptability to satisfy heterogeneous demand from customers. By freeing themselves from the constraints of restricted funding, however, they enter a market place of risk and competition with for-profit enterprises. Their distinctive value lies in social and/or environmental outcomes, which largely remain undocumented.

Developments in accounting for social value will enable social enterprises to accurately measure their performance and prove their worth to stakeholders and policy makers. The desire to achieve measurable outcomes

should, however, resist the pressure to only deliver those services where outcomes can be achieved and measured. Social enterprises have a valuable role to play in creating and sustaining civil society, the benefits of which are consumed by the whole of society. It is in society's interest therefore to ensure that their contribution in this respect is protected.

## REFERENCES

- Alexander, J. (2000). Adaptive Strategies of Non-Profit Human Service Organizations in an Era of Devolution and New Public Management. *Non-Profit Management and Leadership* 10: 287–303.
- Andreason, A. R. (1996). Profits for Non-Profits: Find a Corporate Partner. *Harvard Business Review* 74: 47–59.
- Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: a Theory of Warm Glow Giving. *Economic Journal* 100: 464–477.
- Austin, J. (2000). *The Collaboration Challenge: How Non-Profits and Businesses Succeed Through Strategic Alliances*. Jossey Bass, San Francisco.
- Badelt, C. H. (1997). Entrepreneurship Theories of the Non-Profit Sector. *Voluntas* 8 (2): 162–178.
- Bank of England (2003). *The Financing of Social Enterprises: A Special Report by the Bank of England*; [www.bankofengland.co.uk](http://www.bankofengland.co.uk).
- Ben-Ner, A. and Van Hoomissen, T. (1991). Non-Profit Organizations in the Mixed Economy, a Demand and Supply Analysis. *Annals of Public and Cooperative Economics* 62 (3): 519–550.
- Bilodeau, M. and Slivinski, A. (1998). Rational Non-Profit Entrepreneurship. *Journal of Economics and Management Strategy* 7: 551–571.
- Bornstein, D. (2004). *How to Change the World. Social Entrepreneurs and the Power of New Ideas*. Oxford University Press, Oxford.
- Bornstein, D. (1998). Changing the World on a Shoestring: an Ambitious Foundation Promotes Social Change by Finding 'Social Entrepreneurs'. *The Atlantic Monthly* 281 (1): 34–39.
- Borzaga, C. and Defourny, J. (Eds) (2001). *The Emergence of Social Enterprise*. Routledge, London.
- Boschee, J. and McClurg, J. (2003). Toward a Better Understanding of Social Entrepreneurship: Some Important Definitions. Unpublished paper, [http://www.se-alliance.org/better\\_understanding.pdf](http://www.se-alliance.org/better_understanding.pdf).
- Boschee, J. (1998). *What Does it Take To Be a Social Entrepreneur?* [www.socialentrepreneurs.org/whatdoes.html](http://www.socialentrepreneurs.org/whatdoes.html).
- Brinckerhoff, P. C. (2000). *Social Entrepreneurship: The Art of Mission-Based Venture Development*. Jossey Bass, San Francisco.
- Campbell, D. (2002). Outcomes Assessment and the Paradox of Non-Profit Accountability. *Non-Profit Management and Leadership* 12 (3): 243–259.
- Casson, M. (1999). *The Entrepreneur: An Economic Theory*. Edward Elgar, Oxford.
- Catford, J. (1998). Social Entrepreneurs are Vital for Health Promotion—But They Need Supportive Environments Too. *Health Promotion International* 13 (2): 95–97.
- Chang, C. and Tuckman, H. (1990). Do Higher Property Tax Rates Increase the Market Share of Non-Profit Hospitals? *National Tax Journal* 43: 175–188.
- Chell, E., Hayworth, J. and Breatly, S. (1991). *The Entrepreneurial Personality: Concepts, Cases and Categories*. Routledge, London.

- Clarke, L. and Estes, C. (1992). Sociological and Economic Theories of Markets and Non-Profit: Evidence from Home Health Organizations. *American Journal of Sociology* 97 (4): 945–969.
- Community Wealth Ventures (2003). *Powering Social Change. Lessons on Community Wealth Generation for Non-Profit Sustainability*; [www.communitywealth.com](http://www.communitywealth.com).
- Cordes, J. J., Steuerle, C. E. and Twombly, E. T. (2004). Dimensions of Non-Profit Entrepreneurship: an Exploratory Essay. In *Public Policy and the Economics of Entrepreneurship*, D. Holtz-Eakin and H. S. Rosen (Eds), pp. 115–152. MIT Press, Cambridge.
- Dart, R. (2004). The Legitimacy of Social Enterprise. *Non-Profit Management and Leadership* 14 (4): 411–425.
- Dees, J. G. (1998). Enterprising Non-Profits: What Do You Do When Traditional Sources of Funding Fall Short? *Harvard Business Review* 76: 55–67.
- Dees, J. G. (1996). *The Social Enterprise Spectrum: Philanthropy to Commerce*. Harvard Business School Cases. Case no 9-396-343. Harvard Business School, Boston, MA.
- Dees, J. G., Emerson, J. and Economy, P. (2002). *Strategic Tools for Social Entrepreneurs. Enhancing the Performance of your Enterprising Non-Profit*. John Wiley, New York.
- Dees, J. G., Emerson, J. and Economy, P. (2001). *Enterprising Non-Profits: A Toolkit for Social Entrepreneurs*. Wiley & Sons, New York.
- Defourny, J. (2001). Introduction: From Third Sector to Social Enterprise. In *The Emergence of Social Enterprise*, C. Borzaga and J. Defourny (Eds), pp. 1–28. Routledge, London.
- DiMaggio, P. J. and Anheier, H. K. (1990). The Sociology of Non-Profit Organizations and Sectors. *Annual Review of Sociology* 16: 137–159.
- Drucker, P. (1990). *Managing the Non-Profit Organization: Practices and Principles*. Harper Collins, New York.
- Drucker, P. (1985). *Innovation and Entrepreneurship*. Harper Row, New York.
- DTA (2003). *Fabulous Beasts*. Development Trusts Association; [www.dta.org.uk](http://www.dta.org.uk).
- DTI (2002). *Social Enterprise—a Strategy for Success*. The Department of Trade and Industry; [www.dti.gov.uk](http://www.dti.gov.uk).
- Easley, D. and O'Hara, M. (1983). The Economic Role of Non-Profit Firms. *Bell Journal of Economics* 14: 531–538.
- Eikenberry, A. M. and Kluver, J. D. (2004). The Marketization of the Non-Profit Sector: Civil Society at Risk. *Public Administration Review* 64: 132–140.
- Elliott, B., Katsioloudes, M. and Weldon R. (1998). Non-Profit Organizations and the Internet. *Non-Profit Management and Leadership* 8 (3): 297–303.
- Emerson, J. (2001). *Social Intent and Returns: An Exploration of Investors, Instruments and the Pursuit of a Social Return on Investment*. The Roberts Enterprise Development Fund, San Francisco.
- Emerson, J. and Twersky, F. (Eds) (1996). *New Social Entrepreneurs: The Success, Challenge and Lessons of Non-profit Enterprise Creation*. Roberts Foundation, San Francisco.
- Emerson, J. and Wachowicz, J. (1999). *Social Return on Investment: Exploring Aspects of Value Creation in the Non-Profit Sector*. The Roberts Enterprise Development Fund, San Francisco.
- Enjolras, B. (2002). Does the Commercialisation of Voluntary Organizations 'Crowd Out' Voluntary Work? *Annals of Public and Cooperative Economics* 73 (2): 375–398.
- Etherington, S. (2004). Public Services and the Future of the U.K. Voluntary Sector. *International Journal of Non-Profit and Voluntary Sector Marketing* 9: 105–119.
- File, K. M. and Prince, R. A. (1995). Cause-Related Marketing, Philanthropy and the Arts. *Non-Profit Management and Leadership* 5: 249–260.
- Francois, P. (2003). Not-For-Profit Provision of Public Services. *Economic Journal* 113: C53–C62.
- Frumkin, P. (2002). *On Being a Non-Profit, A Conceptual and Policy Primer*. Harvard University Press, Cambridge, MA.

- Frumkin, P. and Andre-Clark, A. (2000). When Missions, Markets and Politics Collide: Values and Strategies in the Non-Profit Human Services. *Non-Profit and Voluntary Sector Quarterly* 29: 141–163.
- Gassler, R. (1990). Non-Profit and Voluntary Economics: a Critical Survey. *Non-Profit and Voluntary Sector Quarterly* 19: 137–149.
- Giloth, R. (1988). Community Economic Development: Strategies and Practices of the 1980's. *Economic Development Quarterly* 2: 343–350.
- Glaeser, E. L. and Shleifer, A. (2001). Not-For-Profit Entrepreneurs. *Journal of Public Economics* 81: 99–115.
- Glaeser, E. L. and Shleifer, A. (1998). *Not-For-Profit Entrepreneurs*. NBER Working Paper 8610, Cambridge, MA.
- Goerke, J. (2003). Taking the Quantum Leap: Non-Profits are Now in Business. An Australian Perspective. *International Journal of Non-Profit and Voluntary Sector Marketing* 8: 317–327.
- Gov (2002). *Private Action Public Benefit: A Review of Charities and the Wider Not-for-Profit Sector*, The Strategy Unit; <http://www.number-10.gov.uk/su/voluntary/report/pdf.htm>.
- Gui, B. (1991). The Economic Rationale of the Non-Profit Sector. *Annals of Public and Cooperative Economics* 62: 551–572.
- Hansmann, H. B. (1980). The Role of Non-Profit Enterprise. *The Yale Law Journal* 89: 835–901.
- Harding, R. and Cowling, M. (2004). *Social Entrepreneurship Monitor United Kingdom*, London Business School; [www.gem.org](http://www.gem.org).
- Herzlinger, R. E. and Kraker, W. S. (1987). Who Profits From Non-Profits? *Harvard Business Review*, January–February: 93–106.
- Hibbert, S. A., Hogg, G. and Quinn, T. (2002). Consumer Response to Social Entrepreneurship: the Case of The Big Issue in Scotland. *International Journal of Non-Profit and Voluntary Sector Marketing* 7 (3): 288–301.
- Holtmann, A. (1988). Theories of Non-Profit Organizations. *Journal of Economic Surveys* 2 (1): 1–45.
- Hornaday, R. W. (1992). Thinking about Entrepreneurship: a Fuzzy Set Approach. *Journal of Small Business Management* 30 (4): 12–23.
- Huxham, C. (1996). *Creating Collaborative Advantage*. Sage, London.
- James, E. (1990). Economic Theories of the Non-Profit Sector. In *The Third Sector*, H. K. Anheier and W. Seibel (Eds), pp. 21–29. De Gruyter, Berlin, New York.
- James, E. (Ed.) (1989). *The Non-Profit Sector in International Perspective*. Oxford University Press, New York.
- James, E. (1987). The Private Non-Profit Provision of Education: a Theoretical Model and Application to Japan. *Journal of Comparative Economics* 10: 255–276.
- Jas, P., Wilding, K., Wainwright, S., Passey, A. and Hems, L. (2002). *The U.K. Voluntary Sector Almanac, 2002*. NCVO, London.
- Jegen, D. L. (1998). Community Development Venture Capital: Creating a Viable Business Model for the Future. *Non-Profit Management and Leadership* 9 (2): 187–200.
- Johnson, S. (2003). Social Entrepreneurship Literature Review. *New Academy Review* 2: 42–56.
- Kagan, S. L. (1991). Examining Profit and Non-Profit Childcare: an Odyssey of Quality and Auspices. *Journal of Social Issues* 47: 87–104.
- Kaplan, R. S. (2001). Strategic Performance Measurement and Management in Non-Profit Organizations. *Non-Profit Management and Leadership* 11 (3): 353–370.
- Kaplan, R. S. and Norton, D. P. (1996). *The Balanced Scorecard: Translating Strategy into Action*. Harvard University Press, Boston.
- Kingston, J. and Bolton, M. (2004). New Approaches to Funding Not-For-Profit Organizations. *International Journal of Non-Profit and Voluntary Sector Marketing* 9 (2): 112–121.

- Kirzner, I. (1997). Entrepreneurial Discovery and the Competitive Market Process; an Austrian Approach. *Journal of Economic Literature* 35 (1): 60–85.
- Knapp, M. (1989). Intersectoral Differences in Cost-Effectiveness: Residential Child Care in England and Wales. In *The Non-Profit Sector in International Perspective*, E. James (Ed.), pp. 193–216. Oxford University Press, New York.
- Kramer, R. M. (1987). Voluntary Organizations and the Personal Social Services. In *The Non-Profit Sector, A Research Handbook*, W. W. Powell (Ed.), pp. 240–257. Yale University Press, New Haven.
- Leadbeater, C. (2004). *The Co-Creation of Social Value*. A workshop held at The Judge Institute of Management, University of Cambridge.
- Leadbeater, C. (1997). *The Rise of the Social Entrepreneur*. Demon, London.
- Letts, C., Ryan, W. P. and Grossman, A. (1997). Virtuous Capital: What Foundations can Learn from Venture Capitalists, *Harvard Business Review* 75 (2): 36–43.
- Mertens, S. (1999). Non-Profit Organizations and the Social Economy: Two Ways of Understanding the Third Sector. *Annals of Public and Cooperative Economics* 70 (3): 501–520.
- Morley, E., Vinson, E. and Hatry, H. P. (2001). *Look at Outcome Measurement in Non-Profit Agencies*, Independent Sector and Urban Institute; [www.urban.org](http://www.urban.org).
- Mort, G. S., Weerawardena, J. and Carnegie, K. (2003). Social Entrepreneurship: Towards Conceptualisation. *International Journal of Non-Profit and Voluntary Sector Marketing* 8 (3): 76–88.
- Moss Kanter, R. (1999). From Spare Change to Real Change: the Social Sector as a Beat Site for Business Innovation. *Harvard Business Review* 77: 122–133.
- Moss Kanter, R. and Summers, D. V. (1987). Doing Well by Doing Good: Dilemmas in Non-Profit Organizations and the Need for a Multi-Constituency Approach. In *The Non-Profit Sector: A Research Handbook*, W. W. Powell (Ed.), pp. 154–166. Yale University Press, New Haven.
- NCVO (2002). *The U.K. Voluntary Sector Almanac 2002*. NCVO, London.
- New Economics Foundation (2004). *Social Return on Investment, Valuing What Matters*; [www.neweconomics.org](http://www.neweconomics.org).
- Oberfield, A. and Dees, G. (1994). Starting a Non-Profit Venture. In *The Entrepreneurial Venture*, W. H. Sahlman, H. H. Stevenson, M. J. Roberts and A. Bhidé (Eds), pp. 566–579. Harvard Business School Press, Boston.
- OECD (2003). *The Non-Profit Sector in a Changing Economy*. OECD, Paris.
- OECD (1999). *Social Enterprises*. OECD, Paris.
- Parker, S. C. (2004). *A Theory of Social Entrepreneurship*. Mimeo, Durham University, England.
- Pearce, J. (2003). *Social Enterprise in Anytown*. Calouste Gulbenkian Foundation, London.
- Pearce, J., Raynard, P. and Zadek, S. (2000). *Social Auditing for Small Organizations*, New Economics Foundation; [www.neweconomics.org](http://www.neweconomics.org).
- Pearce, J. (1994). Enterprise with a Social Purpose. *Town and Country Planning* 63 (March): 84–85.
- Piliavin, J. A. and Charng, H. W. (1990). Altruism: a Review of Recent Theory and Research. *Annual Review of Sociology* 16: 27–65.
- Pralahad, C. K. and Ramaswamy, V. (2004). *The Future of Competition*. Harvard Business School Press, Boston.
- Preston, A. (1993). The Market for Human Resources, Comparing Professional Career Paths in the Public, Private and Non-Profit Sectors. In *Non-Profit Organizations in a Market Economy*, D. Hammack and D. Young (Eds), pp. 177–202. Jossey Bass, San Francisco.
- Preston, A. (1989). The Non-Profit Worker in a For-Profit World. *Journal of Labour Economics* 7 (4): 438–463.
- Preston, A. (1988). The Non-Profit Firm: a Potential Solution to Inherent Market Failures. *Economic Inquiry* 26 (3): 493–506.

- Putnam, R. D. (2000). *Bowling Alone, The Collapse and Revival of American Community*. Touchstone, New York.
- Quarter, J. and Richmond, B. J. (2001). Accounting for Social Value in Non-Profits and For-Profits. *Non-Profit Management and Leadership* 12 (1): 75–85.
- Quarter, J., Mook, L. and Richmond, B. J. (2003). *What Counts: Social Accounting for Non-Profits and Cooperatives*. Prentice Hall, Englewood Cliffs, New Jersey.
- REDF (2001). SROI Methodology; [www.redf.org/methodology](http://www.redf.org/methodology).
- REDF (2000). SROI Analysis Rubicon Bakery; [www.redf.org](http://www.redf.org).
- Reis, T. K. and Clohesy, S. J. (2001). Unleashing New Resources and Entrepreneurship for the Common Good: a Philanthropic Renaissance. *New Directions in Philanthropic Fundraising* 32 (Summer): 109–143.
- Reisman, A. (1991). Enhancing Non-Profit Resources Through Barter. *Non-Profit Management and Leadership* 1: 253–265.
- Richmond, B. J., Mook, L. and Quarter, J. (2003). Social Accounting for Non-Profits, Two Models. *Non-Profit Management and Leadership* 13: 308–324.
- Rose-Ackerman, S. (1997). Altruism, Ideological Entrepreneurs and the Non-Profit Firm. *Voluntas* 8: 120–134.
- Rose-Ackerman, S. (1996). Altruism, Non-Profits and Economic Theory. *Journal of Economic Literature* 34: 701–728.
- Ryan, W. (1999). The New Landscape for Non-Profits. *Harvard Business Review*, January–February: 128–136.
- Salamon, L. M. (1993). The Marketisation of Welfare: Changing Non-Profit and For-Profit Roles in the American Welfare State. *Social Service Review* 67: 16–39.
- Salamon, L. M. and Anheier, H. (1997). *Defining the Non-Profit Sector: a Cross-National Analysis*. Johns Hopkins Non-Profit Sector Series, Manchester University Press, Manchester.
- Salamon, L., Anheier, H. K., List, R., Toepler, S., Sokolowski, S. W. and Associates (1999). *Global Civil Society: Dimensions of the Non-Profit Sector*. John Hopkins Centre for Civil Society Studies, Baltimore.
- Salamon, L., Sokolowski, S. W. and List, R. (2003). *Global Civil Society An Overview*, The Johns Hopkins Comparative Non-Profit Sector Project; [www.jhu.edu](http://www.jhu.edu).
- Salamon, L., Sokolowski, S. W. and Associates (2004). *Global Civil Society: Dimensions of the Nonprofit Sector*. Volume 2. Kumarian Press, Bloomfield, CT.
- Sawhill, J. and Williamson, D. (2001). Mission Impossible: Measuring Success in Non-Profit Organizations. *Non-Profit Management and Leadership* 11 (3): 371–386.
- SBS (2005). *A Survey of Social Enterprises Across the U.K.* Small Business Service, The Department of Trade and Industry; [www.dti.gov.uk](http://www.dti.gov.uk).
- Schiff, J. and Weisbrod, B. (1991). Competition Between For-Profit and Non-Profit Organizations in Commercial Markets. *Annals of Public and Cooperative Economics* 62 (4): 619–639.
- Schneider, J. (2003). Small, Minority-Based Non-Profits in the Information Age. *Non-Profit Management and Leadership* 13: 383–399.
- Schumpeter, J. A. (1934). The Theory of Economic Development. Reproduced in *Entrepreneurship*, 1990, M. Casson (Ed.), pp. 105–134. Edward Elgar, Aldershot.
- SEC (2003). *There's More To Business Than You Think. A Guide to Social Enterprise*. The Social Enterprise Coalition; [www.socialenterprise.org.uk](http://www.socialenterprise.org.uk).
- Segal, L. and Weisbrod, B. (1998). Interdependence of Commercial and Donative Revenues. In *To Profit or Not to Profit: The Commercial Transformation of the Non-Profit Sector*, B. Weisbrod (Ed.), pp. 105–127. Cambridge University Press, Cambridge.
- Shaver, K. G. and Scott, L. R. (1991). Person, Process, Choice: the Psychology of New Venture Creation. *Entrepreneurship Theory and Practice* 16 (2): 23–45.

- Shaw, M. M. (2003). Successful Collaboration Between the Non-Profit and Public Sectors. *Non-Profit Management and Leadership* 14 (1): 107–120.
- Shaw, E., Shaw, J. and Wilson, M. (2001). *Unsung Entrepreneurs: Entrepreneurship for Social Gain*. University of Durham Business School, Durham.
- Skloot, E. (2000). Evolution or Extinction: a Strategy for Non-Profits in the Marketplace. *Non-Profit and Voluntary Sector Quarterly* 29: 315–324.
- Skloot, E. (1987). Enterprise and Commerce in Non-Profit Organizations. In *The Non-Profit Sector: A Research Handbook*, W. W. Powell (Ed.), pp. 380–393. Yale University Press, New Haven.
- Sloan, F. A. (1998). Commercialization in Non-Profit Hospitals. *Journal of Policy Analysis and Management* 17 (2): 234–252.
- Speckbacher, G. (2003). The Economics of Performance Measurement in Non-Profit Organizations. *Non-Profit Management and Leadership* 13 (3): 267–281.
- Steinberg, R. (1993). Public Policy and the Performance of Non-Profit Organizations: a General Framework. *Non-Profit and Voluntary Sector Quarterly* 22: 13–31.
- Sullivan, G. M., Weerawardena, J. and Carnegie, K. (2003). Social Entrepreneurship: Towards Conceptualization. *International Journal of Non-Profit and Voluntary Sector Marketing* 8: 76–88.
- Taylor, B. E., Chait, R. P. and Holland, T. P. (1996). The New Work of the Non-Profit Board. *Harvard Business Review* 74: 36–44.
- Thompson, J. G., Alvy, G. and Lees A. (2000). Social Entrepreneurship—a New Look at the People and the Potential. *Management Decision* 38: 328–338.
- Tuckman, H. P. (1998). Competition, Commercialization and the Evolution of Non-Profit Organizational Structures. *Journal of Policy Analysis and Management* 17 (2): 175–194.
- Varadarajan, P. and Menon A. (1988). Cause-Related Marketing: A Coalignment of Marketing Strategy and Corporate Philanthropy. *Journal of Marketing* 52 (3): 58–74.
- Venkataraman, S. (1997). The Distinctive Domain of Entrepreneurship Research: an Editor's Perspective. In *Advances in Entrepreneurship, Firm Emergence and Growth*, J. Katz and J. Brockhaus (Eds), pp. 119–138. JAI Press, Greenwich, CT.
- Young, D. (1997). Non-Profit Entrepreneurship. *International Encyclopedia of Public Policy and Administration* 1506–1509.
- Young, D. (1986). Entrepreneurship and the Behaviour of Non-Profit Organizations: Elements of a Theory. In *The Economics of Non-Profit Institutions: Studies in Structure and Policy*, S. Rose-Ackerman (Ed.), pp. 161–184. Oxford University Press, New York.
- Young, D. (1983). *If Not for Profit, For What?* D. C. Heath, Lexington, MA.
- Waddock, S. A. and Post, J. E. (1991). Social Entrepreneurs and Catalytic Change. *Public Administration Review* 51 (5): 393–401.
- Wallace, S. L. (1999). Social Entrepreneurship: the Role of Social Purpose Enterprises in Facilitating Community Economic Development. *Journal of Developmental Entrepreneurship* 4: 153–174.
- Weisbrod, B. A. (2004). The Pitfalls of Non-Profits. *Stanford Social Innovation Review* Winter: 40–47.
- Weisbrod, B. A. (Ed.) (1998a). *To Profit or Not to Profit: The Commercial Transformation of the Non-Profit Sector*. Cambridge University Press, Cambridge.
- Weisbrod, B. A. (1998b). The Non-Profit Mission and its Financing. *Journal of Policy Analysis and Management* 17: 165–174.
- Weisbrod, B. A. (1989). Rewarding Performance that is Hard to Measure: the Private Non-Profit Sector. *Science* 244 (4904): 541–546.
- Weisbrod, B. A. (1988). *The Non-Profit Economy*. Harvard University Press, Cambridge, MA.

- Wolfe, A. (1998). What is Altruism? In *Private Action and the Public Good*, W. W. Powell and E. S. Clemens (Eds), pp. 36–46. Yale University Press, New Haven.
- Zeitlow, J. T. (2001). Social Entrepreneurship: Managerial, Finance and Marketing Aspects. *Journal of Non-Profit and Public Sector Marketing* 9 (1–2): 19–43.
- Zimmerman H. (1999). Innovation in Non-Profit Organizations. *Annals of Public and Cooperative Economics* 70 (4): 589–619.



ROBERT W. FAIRLIE

*Department of Economics, University of California, Santa Cruz, CA 95060, USA*

*E-mail: rfairlie@ucsc.edu*

## 15. Entrepreneurship among Disadvantaged Groups: Women, Minorities and the Less Educated

### 1. INTRODUCTION

Policies to promote entrepreneurship and business ownership among disadvantaged groups are widespread. In the United States, for example, there exist at least 650 nonprofit programs providing loans, training and/or technical assistance to disadvantaged entrepreneurs (Aspen Institute, 2002). Many countries have programs providing financial and other assistance to the unemployed to start businesses.<sup>1</sup> Several states in the United States also have programs providing transfers to unemployment insurance recipients and programs promoting self-employment as a way to leave the welfare rolls (Vroman, 1997; Kosanovich, et al. 2001; Guy, Doolittle and Fink, 1991; Raheim, 1997). The federal government and several states have also promoted self-employment as a way to leave the welfare rolls.

There also exist a large number of federal, state and local government programs providing set-asides and loans to minorities, women and other disadvantaged groups.<sup>2</sup> These affirmative action programs, which target government contracts for disadvantaged and minority-owned firms, have been and continue to be extremely controversial. During the late 1970s and 1980s, there was tremendous growth in the value of federal, state and local government contracts reserved for minority-owned businesses in the United States. The purpose of these set-aside programs was to develop minority enterprise, counter the effects of past discrimination and reduce unemployment among minorities. For the last 15 years, however, the state and local programs established in the 1980s have been both judicially and legislatively challenged and in many cases dismantled. The constitutionality of government-sponsored set-aside programs has been

seriously questioned with the 1989 *Richmond v. J.A. Croson Co.* and 1995 *Adarand Constructors Inc. v. Peña* U.S. Supreme Court decisions.

The interest in entrepreneurship and business development programs has been spurred by arguments from academicians and policymakers that entrepreneurship provides a route out of poverty and an alternative to unemployment or discrimination in the labor market.<sup>3</sup> For example, Glazer and Moynihan (1970, p. 36) argue that “business is in America the most effective form of social mobility for those who meet prejudice.” Proponents also note that many disadvantaged groups facing discrimination or blocked opportunities in the wage/salary sector have used business ownership as a source of economic advancement. It has been argued, for example, that the economic success of earlier immigrant groups in the United States, such as the Chinese, Japanese, Jews, Italians and Greeks, is in part due to their ownership of small businesses (see Loewen, 1971; Light, 1972; Baron et al., 1975; Bonacich and Modell, 1980). More recently, Koreans have purportedly used business ownership for economic mobility (Min, 1989, 1993). Finally, stimulating business creation in sectors with high growth potential (e.g., construction, wholesale trade and business services) may represent an effective public policy for promoting economic development and job creation in poor neighborhoods (Bates, 1993).

The desire for entrepreneurship is also strong in many countries around the world. When individuals were asked the question whether they would prefer “being an employee or being self-employed,” a large percentage reported “self-employment” (Blanchflower, Oswald and Stutzer, 2001). Slightly more than 70% of respondents in the United States expressed a desire to be self-employed. In many other countries, including Germany, Italy and Canada, for example, more than half of all individuals reported a desire for self-employment.<sup>4</sup> Interest in self-employment is also strong among disadvantaged groups. More than 60% of young women and 75% of young blacks reported being interested in starting their own business (Kourilsky and Walstad, 1998; Walstad and Kourilsky, 1998). Among young men and whites, 72 and 63% of respondents, respectively, were interested in starting a business. Although many disadvantaged individuals may possess a strong desire for entrepreneurship, they ultimately do not operate successful small businesses due to a lack of knowledge of business opportunities, sector-specific human capital and financial capital.

In this study, I examine entrepreneurship patterns among several disadvantaged groups. I first document rates of business ownership using micro data from the U.S. Current Population Survey (CPS) and U.S. Census, and aggregate data from the OECD Labour Force Statistics, Canadian Census and British Census making comparisons to more advantaged groups. New estimates of self-employment rates for several ethnic and racial groups from the Canadian, U.K. and U.S. Censuses are presented. Next, using micro data from the CPS, I explore differences in entry rates into and exit rates out of self-employment

across groups. This dynamic approach is useful for identifying the causes of differences in self-employment rates between groups and changes over time in self-employment rates. In previous research, I used this approach to analyze the causes of black/white differences in business ownership rates in the United States (Fairlie, 1999). I also use a nonlinear decomposition technique to identify the contributions from racial and ethnic differences in education, assets and other factors to gaps in self-employment entry and exit rates.

Several major disadvantaged groups are analyzed in this study—women, blacks, Latinos, Native Americans, immigrants and the less educated. All of these groups have substantially lower earnings than their more advantaged counterparts. Among year-round, full-time workers, women earn only 66% of what men earn (U.S. Bureau of the Census, 2004). Latinos, blacks and Native Americans earn only two-thirds to three-fourths of the earnings of white, non-Latinos.<sup>5</sup> Immigrants earn 90% of natives, and high school dropouts earn only 43% of college-educated workers. Most of these groups have also been targeted by set-aside programs in the United States. As discussed below, extensive literatures on female, minority and immigrant entrepreneurship currently exist and a relatively small literature is emerging on less-educated entrepreneurs. The lack of research on less-educated workers is somewhat surprising as this group also faces limited opportunities in the wage/salary sector and has experienced declining wages relative to the wages of their college-educated counterparts.<sup>6</sup> Furthermore, similar to other disadvantaged groups limited access to capital may represent a significant barrier to entrepreneurial success for this group.

## 2. BUSINESS OWNERSHIP AMONG DISADVANTAGED GROUPS

Using data from the Current Population Survey (CPS) Annual Demographic Files (ADF), I first examine patterns of self-employment across several disadvantaged groups. These surveys, conducted annually by the U.S. Bureau of the Census and the Bureau of Labor Statistics, are representative of the entire U.S. population and contain observations for more than 130,000 people. Several recent years of the ADFs are combined to increase the precision of estimates, especially for smaller groups.

Self-employed workers are defined as those individuals who identify themselves as self-employed in their own not incorporated or incorporated business on the class of worker question.<sup>7</sup> The class of worker question refers to the job with the most hours during the reference week. I restrict the sample to include only prime-age individuals (ages 25 to 55) to lessen concerns regarding schooling and retirement decisions.

Table 15-1 reports estimates of self-employment ratios and rates by sex, race/ethnicity, immigrant status and education level. The self-employment

TABLE 15-1 *Self-employment rates by gender, race, immigration and education*

| Group                | Self-employment ratio (%) | <i>N</i> | Self-employment rate (%) | <i>N</i> |
|----------------------|---------------------------|----------|--------------------------|----------|
| Men                  | 12.6                      | 49,299   | 14.2                     | 44,099   |
| Women                | 6.6                       | 54,065   | 8.7                      | 40,967   |
| White, non-Latino    | 10.7                      | 80,383   | 12.8                     | 67,448   |
| Black                | 3.8                       | 9216     | 5.0                      | 7112     |
| Latino               | 5.5                       | 8661     | 7.3                      | 6501     |
| Native American      | 4.8                       | 1202     | 6.5                      | 857      |
| Asian                | 8.9                       | 3902     | 11.2                     | 3148     |
| Native-born          | 9.6                       | 91,869   | 11.6                     | 76,226   |
| Immigrant            | 8.9                       | 11,495   | 11.6                     | 8840     |
| High school dropout  | 6.5                       | 10,293   | 10.3                     | 6600     |
| High school graduate | 9.1                       | 33,699   | 11.4                     | 27,010   |
| Some college         | 9.4                       | 28,721   | 11.2                     | 24,117   |
| College              | 11.0                      | 30,651   | 12.4                     | 27,339   |

Notes: (1) The sample consists of individuals ages 25–55. (2) The self-employment ratio is the number of self-employed business owners divided by the population, and the self-employment rate is the number of self-employed business owners working 15 or more hours divided by all workers with 15 or more hours. (3) All estimates are calculated using sample weights provided by the CPS. Source: Current Population Survey, Matched Annual Demographics Surveys (1998–2002).

ratio is defined as the percentage of all individuals ages 25–55 who report being a self-employed business owner. The self-employment rate conditions on employment. For both measures, women have substantially lower levels of self-employment than men. Although female self-employment rates have risen dramatically in recent decades (see, e.g., Aronson, 1991; Devine, 1994a; U.S. Small Business Administration, 1998; Fairlie, 2004), the prevalence of business ownership among women is only 50–60% of that for men.

The low rate of self-employment among women permeates across ethnic/racial groups and countries. Estimates from the 1990 Census indicate that female self-employment rates are typically around 55% of male rates within detailed ethnic/racial groups and rarely deviate from this ratio (Fairlie and Meyer, 1996). In fact, of the 60 detailed ethnic/racial groups studied, only four groups have female/male self-employment rate ratios that lie outside the range of 0.35 to 0.75. British data reveal a similar pattern of low rates among women within ethnic groups although the ratios of female to male rates are generally lower than in the United States. Estimates from the Fourth National Survey of Ethnic Minorities indicate that female/male ratios range from 0.22 to 0.63 with the exception of the Chinese ratio of 0.88 (Clark and Drinkwater, 2000).

An examination of aggregate data from the OECD also indicates that women are less likely to be self-employed than men for every reported country.

TABLE 15-2 *Male and female self-employment rates by country (nonagricultural sectors)*

| Country         | Self-employment rate |         | Female/male ratio |
|-----------------|----------------------|---------|-------------------|
|                 | Women (%)            | Men (%) |                   |
| Australia       | 8.8                  | 14.8    | 0.595             |
| Austria         | 5.4                  | 9.7     | 0.563             |
| Belgium         | 9.7                  | 17.3    | 0.561             |
| Canada          | 7.9                  | 9.5     | 0.831             |
| Czech Republic  | 9.6                  | 20.3    | 0.476             |
| Germany         | 6.2                  | 12.2    | 0.510             |
| Denmark         | 3.7                  | 10.4    | 0.351             |
| Spain           | 11.3                 | 18.5    | 0.611             |
| Finland         | 6.3                  | 12.3    | 0.511             |
| France          | 4.9                  | 8.3     | 0.595             |
| United Kingdom  | 6.4                  | 14.9    | 0.430             |
| Greece          | 17.5                 | 31.4    | 0.555             |
| Hungary         | 8.1                  | 15.1    | 0.539             |
| Ireland         | 5.8                  | 18.2    | 0.316             |
| Iceland         | 6.8                  | 20.5    | 0.329             |
| Italy           | 15.4                 | 28.0    | 0.550             |
| Japan           | 6.7                  | 10.7    | 0.632             |
| Korea           | 20.4                 | 29.5    | 0.694             |
| Mexico          | 27.4                 | 27.1    | 1.010             |
| Netherlands     | 7.4                  | 11.8    | 0.622             |
| Norway          | 2.9                  | 6.8     | 0.423             |
| New Zealand     | 10.1                 | 20.7    | 0.490             |
| Poland          | 8.3                  | 15.2    | 0.547             |
| Portugal        | 13.3                 | 21.2    | 0.628             |
| Slovak Republic | 4.7                  | 12.1    | 0.388             |
| Sweden          | 4.5                  | 12.3    | 0.370             |
| Turkey          | 8.9                  | 27.1    | 0.330             |
| United States   | 5.4                  | 7.2     | 0.741             |
| Average         | 9.1                  | 16.5    | 0.543             |

Notes: (1) Data for Austria and Turkey are from 2001, and data for Belgium are from 1999. (2) Australia, Japan, Norway and the United States classify owner-managers of incorporated businesses as employees. Austria, Czech Republic, Iceland, Italy, New Zealand and Portugal have unknown classifications for incorporated business owners. OECD Labour Force Statistics (2002). See OECD (2002) for more details.

Table 15-2 reports estimates of the percent of all workers who are self-employed in all industries and nonagricultural industries for men and women. Although self-employment rates vary substantially across countries, female rates are substantially lower than male rates in almost every reported country. The average and median female/male self-employment rate ratios across all countries are 0.543 and 0.548%, respectively. Several countries have female self-employment rates that are roughly one third male rates, and only Canada, Mexico and the United States have female/male self-employment rate ratios

greater than 0.7. The U.S. female/male self-employment rate ratio also appears to be inflated because incorporated business owners are not included in the OECD estimates for the United States. Estimates from the CPS reported in Table 15-1 indicate a U.S. female/male self-employment rate ratio of 0.613. Clearly, women are substantially less likely to be self-employed than are men, which is quite consistent around the world and across different ethnic/racial groups.<sup>8</sup>

Returning to estimates from the CPS reported in Table 15-1, a clear ordering of self-employment propensities across ethnic and racial groups emerges. White, non-Latinos and Asians have the highest self-employment rates and ratios. Among white, non-Latinos, 10.7% of the population ages 25–55 is self-employed and 12.8% of the workforce is self-employed. The Asian self-employment rate and ratio are slightly lower. Relative to these two groups, blacks, Native Americans and Latinos are much less likely to be self-employed. The likelihood of business ownership among Latinos is only slightly higher than 50% of that for white, non-Latinos. Native Americans have even lower levels of business ownership. Finally, of the five ethnic/racial groups identified in this analysis blacks have the lowest rates of business ownership. For example, the black self-employment ratio of 3.8% is roughly one-third the white self-employment ratio. Similarly, low rates of black business ownership date back to at least 1910 (see Fairlie and Meyer, 2000). Clearly, the three major disadvantaged minority groups in the United States—blacks, Latinos and Native Americans—are substantially underrepresented in business ownership.

The ordering of self-employment rates across ethnic/racial groups is similar to that reported in previous studies using alternative data sources and years. These include, but are not limited to, estimates for some or all groups from the 1980 Census (Borjas, 1986; Borjas and Bronars, 1989; Light and Rosenstein, 1995), the 1990 Census (Fairlie and Meyer, 1996; Razin and Light, 1998), the General Social Survey (Hout and Rosen, 2000), the Panel Study of Income Dynamics (Fairlie, 1999) and the Survey of Income and Program Participation (Meyer, 1990; Bates, 1997).

Using aggregate data from the 2001 Canadian and U.K. Censuses and micro data from the 2000 U.S. Census, I provide new estimates of self-employment rates for several ethnic and racial groups (see Table 15-3).<sup>9</sup> All ethnic/racial groups that are roughly comparable for at least two of the three countries selected. Black self-employment rates are higher in the United Kingdom than in Canada and the United States, but remain relatively low. Even in the United Kingdom, where 8.3% of blacks are self-employed business owners, this represents less than two-thirds the white rate of business ownership. Two additional disadvantaged groups—Latinos and Natives—have similarly low self-employment rates in both Canada and the United States. For example,

TABLE 15-3 *Self-employment rates by race/ethnicity for selected countries, 2000–2001*

|            | Canada                   |                | United Kingdom           |                | United States            |                |
|------------|--------------------------|----------------|--------------------------|----------------|--------------------------|----------------|
|            | Self-employment rate (%) | Workers (000s) | Self-employment rate (%) | Workers (000s) | Self-employment rate (%) | Workers (000s) |
| Total      | 12.0                     | 15,516         | 13.7                     | 22,796         | 10.6                     | 115,146        |
| White      | 12.4                     | 13,208         | 13.6                     | 21,277         | 11.8                     | 85,743         |
| Black      | 6.1                      | 315            | 8.3                      | 424            | 4.8                      | 11,368         |
| Latino     | 7.9                      | 114            |                          |                | 7.2                      | 10,696         |
| Native     | 7.2                      | 377            |                          |                | 7.8                      | 808            |
| Asian      | 11.0                     | 1284           | 18.7                     | 849            | 10.9                     | 4034           |
| Chinese    | 13.3                     | 477            | 25.5                     | 89             | 11.2                     | 984            |
| Indian     | 10.0                     | 374            | 17.0                     | 445            | 10.7                     | 693            |
| Vietnamese | 8.8                      | 74             |                          |                | 11.0                     | 411            |
| Korean     | 32.3                     | 43             |                          |                | 23.8                     | 386            |
| Japanese   | 13.5                     | 36             |                          |                | 11.7                     | 345            |
| Filipino   | 3.6                      | 180            |                          |                | 5.0                      | 820            |

Notes: (1) Estimates are from the Canadian 2001 Census, the U.K. 2001 Census and the U.S. 5% Public Use Microdata Sample from the 2000 Census. (2) Canadian minority groups include multiracial responses to the race question. Canadian whites, and all U.S. and U.K. groups include only mono-racial responses to the race question.

only 7.2% of Latinos are self-employed business owners in the United States and 7.9% of Latinos in Canada are self-employed.

There exists substantial heterogeneity across Asian groups. Only 3.6 and 5.0% of Filipinos are self-employed in Canada and the United States, respectively. In contrast, 32.3% of Koreans are self-employed business owners in Canada and 23.8% of Koreans are self-employed in the United States. Another interesting finding is that Chinese, Indians and all Asians have substantially higher rates of business ownership in the United Kingdom than in Canada and the United States.

The estimates reported in Table 15-3 indicate a clear pattern in ethnic/racial entrepreneurship—disadvantaged groups, such as blacks, Latinos and Natives, have relatively low rates of business ownership in all of the countries reported. Thus, low rates of business ownership among these ethnic/racial groups are not peculiar to the United States or one country. Although more research is needed, disadvantaged groups may have similar characteristics that are associated with low levels of entrepreneurship or face similar institutional barriers such as consumer or lending discrimination in each of the countries.

Another disadvantaged group that has received considerable attention in the literature is immigrants. Although immigrants appear to be disadvantaged along many other lines, such as education, income and wealth, their

propensity to own businesses is comparable to native-born Americans. The self-employment ratio among immigrants is only slightly lower than the native ratio and the immigrant self-employment rate is the same as the native rate.

A few recent studies have focused on an additional disadvantaged group—the less educated (see, e.g., Fairlie, 2004; Krashinsky, 2004). Estimates from the CPS indicate that only 6.5% of individuals who do not have a high school diploma are self-employed. In contrast, 11% of college-educated individuals own a business. The differences are smaller, however, after conditioning on employment. The fraction of less-educated individuals who are not employed is higher than that of college-educated individuals.

Estimates from the CPS indicate that disadvantaged groups generally have low rates of business ownership. Although these groups may face limited opportunities in the wage/salary sector, their rates of business ownership are substantially lower than rates for more advantaged groups even conditioning on employment. The major exception studied here is immigrants who have roughly similar levels of self-employment as native-born Americans.

### 3. THE DYNAMICS OF BUSINESS OWNERSHIP

The large disparities in self-employment ratios and rates noted above are created by group differences in transition rates into and out of self-employment. In fact, the steady-state self-employment rate in a simple model of two labor market states is simply equal to  $E/(E + X)$ , where  $E$  is the entry rate into self-employment and  $X$  is the exit rate from self-employment. In a more complicated model with several possible states, the steady-state self-employment ratio is a function of the transition rates to and from each state and their relative shares of the population. A comparison of self-employment transitions across groups may provide insights into the causes of disparities in business ownership.

In previous research using a two-state model, I found that the low rate of self-employment among blacks is due to a black transition rate into self-employment that is approximately one-half the white rate, and a black transition rate out of self-employment that is twice the white rate (Fairlie, 1999). Building on these results, I examined transition patterns among additional disadvantaged groups allowing for three possible states—not employed, wage/salary employed and self-employed. The use of additional labor market states is becoming increasingly popular in the empirical literature on self-employment (see Constant and Zimmerman, 2004; Martinez-Granado, 2002; Kuhn and Schuetze, 2001; Carrasco, 1999, for a few recent examples). For some disadvantaged groups, unemployment may represent a common path into self-employment.

Although the CPS ADFs are primarily used as cross-sectional samples in the previous literature, one-year transitions can be identified by linking



consecutive surveys. Households in the CPS were interviewed each month over a four-month period. Eight months later they were re-interviewed in each month of a second four-month period. Thus, individuals who were interviewed in March of one year were interviewed again in March of the following year. The rotation pattern of the CPS makes it possible to match the information from one survey to the following survey creating a one-year panel for up to half of all respondents in a given ADF. To match data from one survey to the next, I used the procedure described in Madrian and Lefgren (2000).

Table 15-4 reports estimates of transition matrices by sex, race/ethnicity, immigrant status and education level.<sup>10</sup> The first three columns of the table report the percentage of individuals in a specific labor market state in the first survey year who were not employed, employed in a wage/salary job or self-employed one year later. Women have lower rates of entry into self-employment from both nonemployment and wage/salary employment. For example, only 2.1% of women in wage/salary employment start businesses the following year, whereas 3% of wage/salary men switch to self-employment. Women also have substantially higher exit rates from self-employment than men contributing to their relatively low rate of self-employment. Slightly more than a third of all self-employed women leave by the following year compared to 24.6% of self-employed men. Women are also much more likely to make the transition from self-employment to nonemployment. Thus, it appears as though the lower entry rate into self-employment among women and the higher exit rate contribute to their relatively low rate of business ownership.<sup>11</sup>

Of all reported ethnic/racial groups, Asians have the highest entrepreneurship rates. Three and a half percent of Asian wage/salary workers start businesses which is higher than the white wage/salary to self-employment transition rate of 2.7%. Asians, however, have a lower retention rate in self-employment than whites. The result is roughly similar self-employment rates for Asians and whites.

Blacks, Latinos and Native Americans are less likely to start businesses than are whites. All three groups are also more likely to leave self-employment. The differences in transition probabilities between these disadvantaged groups and whites are striking, especially for blacks. Only 1.2% of wage/salary blacks become entrepreneurs, which is less than half the white rate of entrepreneurship. For all three disadvantaged minority groups, exit rates are at least 40%, whereas the white exit rate is 26.8%. Another interesting finding is that these higher exit rates are partly driven by higher transition probabilities to nonemployment. For all three groups, more than 1 out of 10 business owners is nonemployed the following year.

The results for differences between blacks and whites are roughly consistent with those from the PSID reported in Fairlie (1999). Estimates from the PSID indicate that 2% of black men and 4% of white men enter self-

TABLE 15-4 *Transition matrices by gender, race, immigration and education*

| Total                  |                | Year $t + 1$           |               |                         |        |
|------------------------|----------------|------------------------|---------------|-------------------------|--------|
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 75.8%          | 21.2%                  | 3.0%          | 17.9%                   | 18,298 |
| Wage/salary employment | 6.2%           | 91.3%                  | 2.5%          | 72.5%                   | 74,795 |
| Self-employed          | 6.3%           | 21.8%                  | 71.9%         | 9.5%                    | 10,271 |
| Men                    |                | Year $t + 1$           |               |                         |        |
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 71.5%          | 25.0%                  | 3.6%          | 11.2%                   | 5200   |
| Wage/salary employment | 5.1%           | 91.9%                  | 3.0%          | 76.1%                   | 37,562 |
| Self-employed          | 3.7%           | 20.9%                  | 75.4%         | 12.6%                   | 6537   |
| Women                  |                | Year $t + 1$           |               |                         |        |
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 77.7%          | 19.6%                  | 2.7%          | 24.2%                   | 13,098 |
| Wage/salary employment | 7.2%           | 90.7%                  | 2.1%          | 69.2%                   | 37,233 |
| Self-employed          | 11.0%          | 23.4%                  | 65.6%         | 6.6%                    | 3734   |
| Native-born            |                | Year $t + 1$           |               |                         |        |
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 76.1%          | 20.9%                  | 3.0%          | 17.3%                   | 15,643 |
| Wage/salary employment | 5.9%           | 91.6%                  | 2.5%          | 73.1%                   | 66,921 |
| Self-employed          | 6.1%           | 21.4%                  | 72.5%         | 9.6%                    | 9305   |
| Immigrant              |                | Year $t + 1$           |               |                         |        |
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 73.8%          | 23.5%                  | 2.7%          | 22.9%                   | 2655   |
| Wage/salary employment | 8.2%           | 88.4%                  | 3.4%          | 68.2%                   | 7874   |
| Self-employed          | 7.7%           | 25.3%                  | 67.0%         | 8.9%                    | 966    |
| White                  |                | Year $t + 1$           |               |                         |        |
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total | $N$    |
| Non-employment         | 76.0%          | 20.7%                  | 3.2%          | 16.5%                   | 12,935 |
| Wage/salary employment | 5.6%           | 91.7%                  | 2.7%          | 72.8%                   | 58,409 |
| Self-employed          | 5.7%           | 21.1%                  | 73.2%         | 10.7%                   | 9039   |

TABLE 15-4 *Continued*

| Black                  |                   |                        |               |                              |          |
|------------------------|-------------------|------------------------|---------------|------------------------------|----------|
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 76.9%             | 21.2%                  | 1.9%          | 22.7%                        | 2104     |
| Wage/salary employment | 8.6%              | 90.2%                  | 1.2%          | 73.4%                        | 6760     |
| Self-employed          | 13.3%             | 26.9%                  | 59.8%         | 3.8%                         | 352      |
| Latino                 |                   |                        |               |                              |          |
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 72.5%             | 25.3%                  | 2.3%          | 25.4%                        | 2160     |
| Wage/salary employment | 9.4%              | 88.5%                  | 2.1%          | 69.1%                        | 6014     |
| Self-employed          | 11.0%             | 28.3%                  | 60.7%         | 5.5%                         | 487      |
| Native American        |                   |                        |               |                              |          |
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 78.2%             | 20.4%                  | 1.4%          | 26.5%                        | 345      |
| Wage/salary employment | 6.7%              | 91.3%                  | 2.0%          | 68.7%                        | 797      |
| Self-employed          | 20.9%             | 24.0%                  | 55.0%         | 4.8%                         | 60       |
| Asian                  |                   |                        |               |                              |          |
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 74.2%             | 21.9%                  | 3.9%          | 20.0%                        | 754      |
| Wage/salary employment | 6.8%              | 89.6%                  | 3.5%          | 71.1%                        | 2815     |
| Self-employed          | 5.8%              | 25.1%                  | 69.1%         | 8.9%                         | 333      |
| High school dropout    |                   |                        |               |                              |          |
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 82.1%             | 16.2%                  | 1.6%          | 36.6%                        | 3693     |
| Wage/salary employment | 11.1%             | 86.3%                  | 2.7%          | 56.9%                        | 5928     |
| Self-employed          | 12.4%             | 22.6%                  | 65.0%         | 6.5%                         | 672      |
| High school graduate   |                   |                        |               |                              |          |
|                        | Year <i>t</i> + 1 |                        |               |                              |          |
| Year <i>t</i>          | Non-employment    | Wage/salary employment | Self-employed | Share of year <i>t</i> total | <i>N</i> |
| Non-employment         | 75.6%             | 21.5%                  | 3.0%          | 20.4%                        | 6689     |
| Wage/salary employment | 7.2%              | 90.4%                  | 2.4%          | 70.6%                        | 23,775   |
| Self-employed          | 6.8%              | 22.0%                  | 71.2%         | 9.1%                         | 3235     |

TABLE 15-4 *Continued*

| Some college           |                | Year $t + 1$           |               |                         | $N$    |
|------------------------|----------------|------------------------|---------------|-------------------------|--------|
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total |        |
| Non-employment         | 73.9%          | 22.8%                  | 3.3%          | 16.1%                   | 4604   |
| Wage/salary employment | 6.1%           | 91.6%                  | 2.3%          | 74.5%                   | 21,245 |
| Self-employed          | 5.5%           | 22.4%                  | 72.1%         | 9.4%                    | 2872   |

| College                |                | Year $t + 1$           |               |                         | $N$    |
|------------------------|----------------|------------------------|---------------|-------------------------|--------|
| Year $t$               | Non-employment | Wage/salary employment | Self-employed | Share of year $t$ total |        |
| Non-employment         | 72.0%          | 23.9%                  | 4.1%          | 11.0%                   | 3312   |
| Wage/salary employment | 4.1%           | 93.0%                  | 2.9%          | 77.9%                   | 23,847 |
| Self-employed          | 5.2%           | 21.1%                  | 73.7%         | 11.0%                   | 3492   |

Notes: Current Population Survey, Matched Annual Demographic Surveys (1998–2003). (1) The sample consists of individuals ages 25–55. (2) All estimates are calculated using sample weights provided by the CPS.

employment annually, and 36.6% of black men and 18.5% of white men exit self-employment annually. Excluding women and nonemployment from the CPS sample for comparability, I find black and white transition rates into self-employment of 2.0 and 3.3, respectively. The black exit rate is 29.1% and the white exit rate is 20.8%.

Table 15-4 also reports estimates by immigrant status. Immigrants have a higher transition rate into self-employment from wage/salary employment than natives, but not from nonemployment. The total transition rate into self-employment, however, is higher among immigrants. On the other hand, immigrants are more likely than natives to leave self-employment. Thirty-three percent of self-employed immigrants leave annually, whereas 27.5% of self-employed natives leave annually.

Estimates from transition matrices by major education level also reveal a few interesting patterns. First, education and entry into self-employment from wage/salary employment have an U-shaped relationship. Entry into business ownership from nonemployment, however, is clearly increasing with education. Only 1.6% of nonemployed high school dropouts start a business the following year compared to 4.1% of nonemployed college-educated individuals.<sup>12</sup> Second, exit rates from self-employment decrease with education contributing to the positive relationship between self-employment and education shown in Table 15-1. Finally, a much higher percentage of those leaving self-employment move to nonemployment among the less educated than among the more educated. In fact, the probability of becoming nonemployed conditional on self-

employment is more than twice as high for high school dropouts as it is for college graduates.

To summarize, all disadvantaged groups, with the exception of immigrants, have relatively low rates of entering self-employment and high rates of exiting self-employment. Disadvantaged groups also generally have high rates of movement from self-employment to nonemployment. Interestingly, however, their rates of entry into self-employment from nonemployment are lower across all of the dimensions analyzed here, including immigrants. This finding contrasts with disadvantaged theory, which states that disadvantages such as poverty, unemployment and discrimination push certain groups into self-employment instead of wage/salary work. Furthermore, although disadvantaged groups have relatively high rates of nonemployment, the preponderance of entrants into self-employment come from wage/salary employment for all groups.

#### 4. THE DETERMINANTS OF SELF-EMPLOYMENT ENTRY AND EXIT

To identify the independent effects of sex, race, immigrant status and education, I estimate logit regressions for self-employment transition probabilities. Separate logit regressions are estimated for the probability of entry into self-employment and the probability of exit from self-employment. To simplify, I do not estimate separate regressions for the probability of entry from nonemployment and wage/salary employment. I also do not distinguish between leaving self-employment for nonemployment or for wage/salary employment. Instead, nonemployment and wage/salary employment are grouped together and a dummy variable is included for nonemployment in the entry regression. The *nonemployed* represent 22.5% of entrants into self-employment and 22.3% of leavers from self-employment. I find that results conditioning on employment in both years, and thus focusing on wage/salary to self-employment and self-employment to wage/salary transitions, are qualitatively similar.<sup>13</sup>

Estimates for the probability of entry into self-employment are reported in Table 15-5 and are discussed first.<sup>14</sup> Marginal effects and their standard errors are reported.<sup>15</sup> All of the independent variables are measured in the first year surveyed, which is prior to when the self-employment entry decision is measured. Specification 1 reports estimates from four separate regressions that include dummies for sex, racial groups, immigrant status and education levels. Each regression includes only one set of variables. The marginal effects estimates on these dummies capture a weighted average of differences in the transition probabilities into self-employment reported in Table 15-4. The comparison or left-out groups in the four separate regressions are men, non-Latino whites, natives and college graduates, respectively. The coefficient

TABLE 15-5 Logit regressions for probability of entry into self-employment

| Explanatory variables | Specification       |                           |                           |                           |                           |             |
|-----------------------|---------------------|---------------------------|---------------------------|---------------------------|---------------------------|-------------|
|                       | (1)                 | (2)                       | (3)                       | (4)                       | (5)                       |             |
| Female                | -0.0092<br>(0.0011) | **<br>-0.0089<br>(0.0011) | **<br>-0.0089<br>(0.0011) | **<br>-0.0104<br>(0.0011) | **<br>-0.0103<br>(0.0011) | **<br>2.576 |
| Black                 | -0.0179<br>(0.0024) | **<br>-0.0170<br>(0.0024) | **<br>-0.0178<br>(0.0026) | **<br>-0.0154<br>(0.0027) | **<br>-0.0139<br>(0.0027) | **<br>1.96  |
| Latino                | -0.0073<br>(0.0021) | **<br>-0.0092<br>(0.0024) | **<br>-0.0040<br>(0.0029) | -0.0048<br>(0.0030)       | -0.0036<br>(0.0030)       |             |
| Native American       | -0.0109<br>(0.0057) | -0.0095<br>(0.0057)       | -0.0091<br>(0.0057)       | -0.0147<br>(0.0057)       | -0.0134<br>(0.0057)       | *           |
| Asian                 | 0.0046<br>(0.0025)  | -0.0012<br>(0.0029)       | -0.0066<br>(0.0058)       | -0.0082<br>(0.0059)       | -0.0072<br>(0.0059)       |             |
| Immigrant             | 0.0035<br>(0.0016)  | *<br>0.0077<br>(0.0021)   | **                        |                           |                           |             |
| White*Immigrant       |                     |                           | 0.0099<br>(0.0027)        | **<br>0.0109<br>(0.0027)  | **<br>0.0116<br>(0.0028)  | **          |
| Black*Immigrant       |                     |                           | 0.0141<br>(0.0064)        | *<br>0.0158<br>(0.0065)   | *<br>0.0164<br>(0.0065)   | *           |
| Latino*Immigrant      |                     |                           | -0.0016<br>(0.0040)       | -0.0019<br>(0.0040)       | -0.0016<br>(0.0040)       |             |
| Asian*Immigrant       |                     |                           | 0.0145<br>(0.0064)        | *<br>0.0157<br>(0.0064)   | *<br>0.0161<br>(0.0064)   | *           |
| High school dropout   | -0.0102<br>(0.0021) | **<br>-0.0082<br>(0.0022) | **<br>-0.0075<br>(0.0022) | **<br>-0.0106<br>(0.0022) | **<br>-0.0078<br>(0.0023) | **          |
| High school graduate  | -0.0059<br>(0.0013) | **<br>-0.0046<br>(0.0013) | **<br>-0.0045<br>(0.0013) | **<br>-0.0058<br>(0.0014) | **<br>-0.0036<br>(0.0014) | *           |
| Some college          | -0.0045<br>(0.0014) | **<br>-0.0030<br>(0.0014) | *<br>-0.0030<br>(0.0014)  | *<br>-0.0043<br>(0.0014)  | **<br>-0.0031<br>(0.0014) | *           |

TABLE 15-5 *Continued*

| Explanatory variables          | Specification |     |     |                     |                          |
|--------------------------------|---------------|-----|-----|---------------------|--------------------------|
|                                | (1)           | (2) | (3) | (4)                 | (5)                      |
| Age                            |               |     |     | 0.0016<br>(0.0007)  | *<br>0.0015<br>(0.0007)  |
| Age squared/100                |               |     |     | -0.0018<br>(0.0009) | *<br>-0.0019<br>(0.0009) |
| Married                        |               |     |     | 0.0060<br>(0.0018)  | **<br>0.0052<br>(0.0019) |
| Previously married             |               |     |     | 0.0034<br>(0.0022)  | 0.0033<br>(0.0022)       |
| Number of children             |               |     |     | -0.0003<br>(0.0010) | -0.0006<br>(0.0010)      |
| Number of children squared     |               |     |     | 0.0005<br>(0.0002)  | **<br>0.0006<br>(0.0002) |
| Not employed                   |               |     |     | 0.0078<br>(0.0013)  | **<br>0.0084<br>(0.0014) |
| Home owner                     |               |     |     |                     | *<br>0.0039<br>(0.0015)  |
| Dividend income (000s)         |               |     |     |                     | 0.0014<br>(0.0009)       |
| Dividend income squared (000s) |               |     |     |                     | -0.0001<br>(0.0001)      |
| Interest income (000s)         |               |     |     |                     | 0.0013<br>(0.0004)       |

TABLE 15-5 *Continued*

| Explanatory variables          | Specification |         |         |         |                          |
|--------------------------------|---------------|---------|---------|---------|--------------------------|
|                                | (1)           | (2)     | (3)     | (4)     | (5)                      |
| Interest income squared (000s) |               |         |         |         | 0.0000                   |
| Rental income (000s)           |               |         |         |         | (0.0000)<br>0.0023<br>** |
| Rental income squared (000s)   |               |         |         |         | (0.0007)<br>0.0000       |
| Mean of dependent variable     | 0.0270        | 0.0270  | 0.0270  | 0.0270  | (0.0000)<br>0.0270       |
| Log likelihood value           |               | -11,461 | -11,457 | -11,366 | -11,189                  |
| Sample size                    | 93,093        | 93,093  | 93,093  | 93,093  | 91,819                   |

Notes: Current Population Survey, Matched Annual Demographic Surveys (1998–2003). (1) The sample consists of individuals (ages 25–55) who are not self-employed business owners in year  $t$ . (2) All independent variables are measured in the first year surveyed. (3) Marginal effects and their standard errors are reported. Statistical significance at the 0.05 and 0.01 levels are denoted by \* and \*\*, respectively. (4) All specifications include a constant, and specifications 3–4 also include dummy variables for Census divisions, central city status and year effects. Specification 1 includes separate regressions for each of the four sets of listed variables.



estimates create a baseline for comparison to other specifications and indicate that disadvantaged groups, with the exception of immigrants, have substantially lower business entry rates than do the comparison advantaged groups.

It is well known that race, immigrant status and education levels are related. To explore the effects of these correlations on self-employment entry rates for disadvantaged groups, Specification 2 includes the sex, race, immigrant status and education dummies in one regression. The marginal effects estimates capture differences in transition probabilities accounting for these correlations. Two key patterns emerge from comparing these estimates to the previous estimates. First, the black, Native American, high school dropout and high school graduate coefficients become smaller in absolute value because of the correlation between these racial groups and low levels of education. Second, the negative coefficient on Latino becomes larger in absolute value and the positive coefficient on Asian essentially disappears at the same time the positive coefficient on immigrant increases substantially. Apparently, the preponderance of immigrants among Asians and Latinos makes the coefficient estimates for these racial/ethnic groups sensitive to controlling for immigrant status. Overall, the logit estimates clearly indicate that disadvantaged groups, with the exception of immigrants, have substantially lower business entry rates than do the comparison advantaged groups.

Returning to the issue of the interrelatedness of race/ethnicity and immigration, I found that 54.7% of Latinos and 73.9% of Asians are immigrants in my sample, and that these two groups comprise roughly two-thirds of all immigrants. Although immigrants from different races and ethnicities share some common attributes such as language barriers and unfamiliarity with U.S. institutions, they may differ substantially along many other dimensions such as reasons for emigrating, home country economic conditions, networks and unobserved skills. Specification 3 addresses this concern by including interactions between immigrant status and race/ethnicity. A clear pattern emerges—white, Asian and black immigrants are substantially more likely to become business owners than their native-born counterparts, whereas Latino immigrants are slightly less likely to enter self-employment than native Latinos (although the difference is not statistically significant).<sup>16</sup> Apparently, immigrant status has an independent effect on self-employment entry, but its effects differ somewhat by race/ethnicity. Thus, I allow for differential effects by race/ethnicity below.

Specification 4 adds controls for age, marital status, number of children, region of the country, central city status, survey year and nonemployment in the regressions. The probability of entering self-employment increases with age (up to age 43), being married, the number of children and nonemployment. Controlling for these characteristics generally does not have a large effect on the majority of the female, race/immigrant and education coefficients. The exceptions are that the black coefficient declined in absolute value, and

the Native American, Asian and high school dropout coefficients increased in absolute value. Clearly, the inclusion of these controls does not “explain away” the general finding of low rates of entry into self-employment by most disadvantaged groups.

The importance of assets has taken center stage in the literature on the determinants of self-employment. Numerous studies using various methodologies, measures of assets and country micro data explore the relationship between assets and self-employment. Several recent studies estimate the relationship by modeling the decision of wage/salary workers or other non-business owners to switch into self-employment over a fixed period of time.<sup>17</sup> These studies generally find that asset levels (e.g., net worth or asset income) measured in one year increase the probability of entering self-employment by the following year suggesting that entrepreneurs may face liquidity constraints.<sup>18</sup> Recent studies also indicate that blacks have substantially lower levels of assets than whites and that these differences contribute to racial differences in business ownership levels (Bates, 1989; Fairlie, 1999; Fairlie and Robb, 2003). Although less is known for other disadvantaged groups, disparities in asset levels may be large and explain why these groups are also less likely to become business owners.

In Specification 5, I add several measures of assets available in the CPS to the logit regression. Home ownership is included as well as dividend, interest and rental income. Investment and rental income are not a direct measure of assets but are roughly proportional to asset levels. These measures are included separately to allow for differential values on the underlying assets and liquidity. All measures of assets are measured prior to the self-employment decision.<sup>19</sup> As expected, home owners are more likely to enter self-employment.<sup>20</sup> In the presence of liquidity constraints, the ability of owners to borrow against the value of their home, such as home equity loans, may make it easier to finance new business ventures. The relationships between the probability of making a transition into self-employment and dividend, interest and rental income are concave and increasing at the means of each measure of asset income. Similar to previous studies, I find that higher levels of assets increase the probability of entry into self-employment.

Controlling for differences in asset levels reduces (in absolute value) the coefficients for most disadvantaged groups. All of the main effects for ethnic and racial groups become smaller in absolute value (or increase) suggesting that assets levels are relatively low among native-born minorities compared to native-born whites and that these low levels of assets limit their opportunities to start businesses. I also find that the coefficient estimates on all of the immigrant/race interactions increase suggesting that controlling for assets explains an additional amount of the difference in entry rates between racial groups and whites among immigrants. For example, the native Latino differential drops from 0.48 percentage points to 0.36 percentage points after controlling for

assets, whereas the immigrant Latino differential drops from 0.67 to 0.51 percentage points. Low levels of assets appear to limit entrepreneurial opportunities among disadvantaged ethnic and racial groups, however, it is difficult to identify the importance of this factor. I explore this question further in Section 5.

The coefficients decline sharply for high school dropouts and high school graduates after the addition of assets. This result suggests that, all else being equal, less-educated individuals have relatively low levels of assets resulting in lower entry rates. Indeed, a direct comparison of asset levels by education level reveals that high school graduates, and especially high school dropouts, have substantially lower levels of assets than do college graduates. For example, only 60.1% of high school dropouts own a house compared to 83.3% of college graduates, and average interest income among dropouts is \$109 compared to \$1190 among college graduates. The presence of liquidity constraints and relatively low levels of assets appears to limit the ability of less-educated workers to start businesses.

Even controlling for differences in asset levels, the individual's education level has a strong positive effect on entry into self-employment. High school dropouts are nearly a full percentage point less likely to enter self-employment, and high school graduates and individuals with some college are slightly more than 0.3 percentage points less likely to enter self-employment than are college graduates. Estimates from several other countries, however, indicate a generally statistically insignificant relationship between education and self-employment entry (see, e.g., Holtz-Eakin and Rosen, 2004 for Germany; Blanchflower and Meyer, 1994 for Australia; Lin, Picot and Compton, 2001 for Canada).<sup>21</sup> On the other hand, evidence from Eastern European transition economies indicates a positive relationship between schooling and transitions into self-employment.

#### *4.1. Female Entrepreneurship*

Interestingly, the female coefficient increases slightly in absolute value after the inclusion of the main controls and does not change after inclusion of assets. The change in the marginal effects estimate from Specification 3 to Specification 4 is primarily due to the higher percentage of women who are not employed and the higher rate of entry from nonemployment than from wage/salary employment into self-employment. This finding suggests that the female/male difference in self-employment entry rates would be slightly larger if not for the initial difference in nonemployment rates. As expected, controlling for other variables has little effect on the female coefficient estimate because men and women have very similar characteristics.<sup>22</sup> Women are much less likely than men to enter self-employment, all else being equal. Similar results are found using micro data from the European Union. Blanchflower (2000, 2004) found large female/male differences in the probability of being self-

employed after including education and other measurable individual characteristics as well as country dummies.

Recent studies focusing on gender differences in self-employment provide some interesting findings but provide only limited direct evidence on the question of what explains the large gender difference in self-employment rates.<sup>23</sup> For example, these studies find that women who are married to self-employed men are more likely to be self-employed or enter self-employment and that the choice of self-employment is partly driven by the desire for flexible schedules and other family-related reasons for women relative to men (Bruce, 1999; Boden, 1996, 1999; Carr, 1996; Devine, 1994b; Lombard, 2001; Lohmann 2001).<sup>24</sup> Gender earnings differentials in the wage/salary sector may contribute, but there is also considerable evidence indicating large female/male earnings differences in the self-employment sector (Aronson, 1991; Devine, 1994b; Hundley, 2000; U.S. Bureau of the Census, 2004). In the end, unobservable factors, such as different preferences, discrimination, and risk aversion, may be responsible for low levels of female entrepreneurship.<sup>25</sup> As noted above, an interesting finding is that a lower percentage of young women than men report a desire for being self-employed in the United States (Kourilsky and Walstad, 1998). Using a combined sample from many countries, Blanchflower, Oswald and Stutzer (2001) also found a lower probability of preferring self-employment among women after controlling for other factors. In both cases, however, the differences are not large and represent roughly 15 percentage points.

#### *4.2. Transitions out of Self-employment*

Logit regressions are also estimated for the probability of exit from self-employment. Estimates are reported in Table 15-6. Specification 1 reports estimates from four separate regressions that include dummies for sex, ethnic/racial groups, immigrant status and education levels. All disadvantaged groups have relatively high exit rates from self-employment. Specification 2 accounts for the correlations between sex, race, immigrant status and education. Again, women, blacks, Latinos, Native Americans and high school dropouts have relatively high exit rates from self-employment. The difference between immigrants and natives, although positive, is now small and statistically insignificant. The coefficients on all of the race/ethnicity, immigrant and education dummies become smaller in absolute value due to the correlation between these factors.

Specification 3 includes race/immigrant status interactions. Interestingly, white immigrants have high exit rates relative to white natives, whereas black and Asian immigrants have lower exit rates than their native counterparts (although the differences are not statistically significant). Finally, there appears

TABLE 15-6 *Logit regressions for probability of exit from self-employment*

| Explanatory variables | Specification      |                          |                          |                          |                          |    |
|-----------------------|--------------------|--------------------------|--------------------------|--------------------------|--------------------------|----|
|                       | (1)                | (2)                      | (3)                      | (4)                      | (5)                      |    |
| Female                | 0.0989<br>(0.0088) | **<br>0.1001<br>(0.0088) | **<br>0.0999<br>(0.0088) | **<br>0.0997<br>(0.0088) | **<br>0.0993<br>(0.0089) | ** |
| Black                 | 0.1439<br>(0.0216) | **<br>0.1320<br>(0.0216) | **<br>0.1453<br>(0.0229) | **<br>0.1010<br>(0.0235) | **<br>0.0967<br>(0.0238) | ** |
| Latino                | 0.1097<br>(0.0189) | **<br>0.0841<br>(0.0210) | **<br>0.0989<br>(0.0276) | **<br>0.0771<br>(0.0278) | **<br>0.0701<br>(0.0281) | *  |
| Native American       | 0.1432<br>(0.0513) | **<br>0.1263<br>(0.0512) | **<br>0.1303<br>(0.0512) | *0.1456<br>(0.0510)      | **<br>0.1448<br>(0.0510) | ** |
| Asian                 | 0.0546<br>(0.0236) | *<br>0.0383<br>(0.0263)  | **<br>0.1118<br>(0.0434) | **<br>0.1075<br>(0.0438) | *<br>0.1052<br>(0.0439)  | *  |
| Immigrant             | 0.0653<br>(0.0142) | **<br>0.0253<br>(0.0171) |                          |                          |                          |    |
| White*Immigrant       |                    |                          | 0.0586<br>(0.0213)       | **<br>0.0399<br>(0.0214) | 0.0380<br>(0.0215)       |    |
| Black*Immigrant       |                    |                          | -0.0708<br>(0.0654)      | -0.0648<br>(0.0651)      | -0.0707<br>(0.0653)      |    |
| Latino*Immigrant      |                    |                          | 0.0013<br>(0.0369)       | -0.0057<br>(0.0369)      | -0.0116<br>(0.0372)      |    |
| Asian*Immigrant       |                    |                          | -0.0730<br>(0.0510)      | -0.0804<br>(0.0507)      | -0.0826<br>(0.0508)      |    |
| High school dropout   | 0.0906<br>(0.0176) | **<br>0.0699<br>(0.0180) | **<br>0.0703<br>(0.0181) | **<br>0.0749<br>(0.0183) | **<br>0.0720<br>(0.0186) | ** |
| High school graduate  | 0.0168<br>(0.0109) | 0.0109<br>(0.0108)       | 0.0107<br>(0.0108)       | 0.0211<br>(0.0110)       | 0.0168<br>(0.0113)       |    |
| Some college          | 0.0084<br>(0.0113) | 0.0016<br>(0.0112)       | 0.0011<br>(0.0112)       | 0.0089<br>(0.0113)       | 0.0076<br>(0.0115)       |    |

\*\* 2.576  
\* 1.96

TABLE 15-6 *Continued*

| Explanatory variables          | Specification |     |     |                     |                     |
|--------------------------------|---------------|-----|-----|---------------------|---------------------|
|                                | (1)           | (2) | (3) | (4)                 | (5)                 |
| Age                            |               |     |     | -0.0260<br>(0.0059) | **<br>(0.0059)      |
| Age squared/100                |               |     |     | 0.0278<br>(0.0071)  | **<br>(0.0072)      |
| Married                        |               |     |     | -0.0323<br>(0.0159) | *<br>(0.0162)       |
| Previously married             |               |     |     | -0.0084<br>(0.0189) | -0.0073<br>(0.0191) |
| Number of children             |               |     |     | -0.0058<br>(0.0088) | -0.0051<br>(0.0089) |
| Number of children squared     |               |     |     | 0.0017<br>(0.0021)  | 0.0016<br>(0.0021)  |
| Home owner                     |               |     |     |                     | **                  |
| Dividend income (000s)         |               |     |     |                     | -0.0563<br>(0.0136) |
| Dividend income squared (000s) |               |     |     |                     | -0.0059<br>(0.0069) |
| Interest income (000s)         |               |     |     |                     | 0.0003<br>(0.0005)  |
| Interest income squared (000s) |               |     |     |                     | 0.0044<br>(0.0033)  |
|                                |               |     |     |                     | -0.0001<br>(0.0001) |

TABLE 15-6 *Continued*

| Explanatory variables        | Specification |        |        |        |                    |
|------------------------------|---------------|--------|--------|--------|--------------------|
|                              | (1)           | (2)    | (3)    | (4)    | (5)                |
| Rental income (000s)         |               |        |        |        | -0.0056            |
| Rental income squared (000s) |               |        |        |        | (0.0050)<br>0.0001 |
| Mean of dependent variable   | 0.2696        | 0.2696 | 0.2696 | 0.2696 | (0.0002)<br>0.2692 |
| Log Likelihood value         |               | -5875  | -5870  | -5798  | -5714              |
| Sample size                  | 10,271        | 10,271 | 10,271 | 10,271 | 10,145             |

Notes: Current Population Survey, Matched Annual Demographic Surveys (1998–2003). (1) The sample consists of individuals (ages 25–55) who are self-employed business owners in year *t*. (2) All independent variables are measured in the first year surveyed. (3) Marginal effects and their standard errors are reported. Statistical significance at the 0.05 and 0.01 levels are denoted by \* and \*\*, respectively. (4) All specifications include a constant, and specifications 3–4 also include dummy variables for Census divisions, central city status and year effects. Specification 1 includes separate regressions for each of the four sets of listed variables.

to be no difference between transition rates out of self-employment between native and immigrant Latinos.

Specification 4 includes controls for individual characteristics. The exit rate decreases with age (until age 47) and being married. The coefficient estimates for blacks, Latinos and white immigrants become notably smaller, whereas the coefficient estimate for Native Americans becomes larger. In the final specification, I include controls for asset levels. As expected, home ownership decreases the probability of exit from self-employment and the asset income measures generally have a negative relationship with the exit probability. Even after controlling for asset levels, most disadvantaged groups are substantially more likely to leave business ownership annually.

Overall, disadvantaged groups have relatively low rates of entry into self-employment and high rates of exit from self-employment. The only exception is immigrants who have a higher rate of entry into self-employment than natives, but this comparison does not hold for all groups. Among Latinos, immigration is not associated with a higher level of entry. These patterns of low entry rates and high exit rates among disadvantaged groups persist even after controlling for the correlated effects of other disadvantages (i.e., race and education), individual characteristics and asset levels. Again, immigrants represent the exception as the immigrant/native difference in exit rates becomes negligible for some groups after controlling for other factors. I now turn to a more detailed analysis of the causes of low rates of self-employment entry and exit among disadvantaged minority groups.

## 5. IDENTIFYING THE CAUSES OF ETHNIC AND RACIAL DIFFERENCES IN ENTRY AND EXIT RATES

The estimates reported in Tables 15-5 and 15-6 indicate that the relatively low rates of entry and high rates of exit from self-employment among disadvantaged minority groups can be explained, in part, by group differences in education, assets and other individual characteristics. The estimates, however, cannot identify the separate contributions from group differences in each of these variables. To explore these issues further, I employ a variant of the familiar technique of decomposing inter-group differences in a dependent variable into those due to different observable characteristics across groups and those due to different “prices” of characteristics of groups (see Blinder 1973 and Oaxaca, 1973). The technique that I describe here takes into account the nonlinearity of the logit regressions discussed above (see Fairlie 1999, 2003, for more details).

For a linear regression, the standard Blinder-Oaxaca decomposition of the white/minority gap in the average value of the dependent variable,  $Y$ , can be



expressed as:

$$\bar{Y}^W - \bar{Y}^M = [(\bar{X}^W - \bar{X}^M)\hat{\beta}^W] + [\bar{X}^M(\hat{\beta}^W - \hat{\beta}^M)], \tag{1}$$

where  $\bar{X}^j$  is a row vector of average values of the independent variables and  $\hat{\beta}^j$  is a vector of coefficient estimates for race  $j$ . For a nonlinear equation, such as  $Y = F(X\hat{\beta})$ , the decomposition can be written as:

$$\begin{aligned} \bar{Y}^W - \bar{Y}^M = & \left[ \sum_{i=1}^{N^W} \frac{F(X_i^W \hat{\beta}^W)}{N^W} - \sum_{i=1}^{N^M} \frac{F(X_i^M \hat{\beta}^W)}{N^M} \right] \\ & + \left[ \sum_{i=1}^{N^M} \frac{F(X_i^M \hat{\beta}^W)}{N^M} - \sum_{i=1}^{N^M} \frac{F(X_i^M \hat{\beta}^M)}{N^M} \right], \end{aligned} \tag{2}$$

where  $N^j$  is the sample size for race  $j$ . This alternative expression for the decomposition is used because  $\bar{Y}$  does not necessarily equal  $F(\bar{X}\hat{\beta})$ . In both (1) and (2), the first term in brackets represents the part of the racial gap that is due to group differences in distributions of  $X$ , and the second term represents the part due to differences in the group processes determining levels of  $Y$ . To calculate the decomposition, I define  $\bar{Y}$  as the self-employment entry or exit rate and  $F$  as the logistic cumulative distribution function.

An equally valid method of calculating the decomposition is to use the minority coefficient estimates,  $\hat{\beta}^M$ , as weights in estimating the contributions from group differences in the independent variables. This alternative method of calculating the decomposition often provides different estimates, which is the familiar index problem with the Blinder-Oaxaca decomposition technique. A third commonly-used alternative is to weight the first term of the decomposition expression using coefficient estimates from a pooled sample of the two groups or all groups (see, e.g., Oaxaca and Ransom, 1994). I follow this approach to calculate the decompositions. In particular, I use coefficient estimates from logit regressions that include pooled samples of all ethnic and racial groups.

The first term in (2) provides an estimate of the contribution of racial differences in the entire set of independent variables to the racial gap, but I am particularly interested in identifying the effects of group differences in specific variables, such as education and asset levels. To identify contributions from these variables an additional calculation is needed. To simplify, assume that  $X$  includes two variables,  $X_1$  and  $X_2$ . Using coefficient estimates from a logit regression for a pooled sample,  $\hat{\beta}^*$ , the independent contribution of  $X_1$  to the

racial gap can then be expressed as:

$$\frac{1}{NM} \sum_{i=1}^{NM} F(\hat{\alpha}^* + X_{1i}^W \hat{\beta}_1^* + X_{2i}^W \hat{\beta}_2^*) - F(\hat{\alpha}^* + X_{1i}^M \hat{\beta}_1^* + X_{2i}^W \hat{\beta}_2^*). \quad (3)$$

Similarly, the contribution of  $X_2$  can be expressed as:

$$\frac{1}{NM} \sum_{i=1}^{NM} F(\hat{\alpha}^* + X_{1i}^M \hat{\beta}_1^* + X_{2i}^W \hat{\beta}_2^*) - F(\hat{\alpha}^* + X_{1i}^M \hat{\beta}_1^* + X_{2i}^M \hat{\beta}_2^*). \quad (4)$$

The contribution of each variable to the gap is thus equal to the change in the average predicted probability from replacing the black distribution with the white distribution of that variable while holding the distributions of the other variable constant.<sup>26</sup> A useful property of this technique is that the sum of the contributions from individual variables will be equal to the total contribution from all of the variables evaluated with the full sample.

Table 15-7 reports estimates from this procedure for decomposing the gap between the native-born white and minority gaps in self-employment entry rates. I report estimates only for those race/immigrant groups that have large enough sample sizes. The individual contributions from racial differences in education, marital status and children, nonemployment, assets, region and central city status are reported. I first describe the results for native-born blacks, which are reported in Specification 1. The native white/black gap in the self-employment entry rate is large (0.0144). Racial differences in sex and age explain virtually none of the gap. Marital status and children explain only a small part of the gap (5%). This contribution is primarily due to blacks having a substantially lower probability of currently being married than whites and the positive effect of marriage on entry into self-employment. Slightly more of the gap is explained by relatively low levels of education among blacks. In the sample, 14.3% of blacks are high school dropouts compared to only 6.2% of whites.

As expected, the largest factor explaining racial disparities in business creation rates are differences in asset levels.<sup>27</sup> Lower levels of assets among blacks account for 15.5% of the white/black gap in the probability of entry into self-employment. In the presence of liquidity constraints, low levels of assets appear to limit opportunities for blacks to start businesses. The finding is very similar to estimates reported in Fairlie (1999) for men in the PSID. Estimates from the PSID indicate that 13.9 to 15.2% of the black/white gap in the transition rate into self-employment can be explained by differences in assets.

TABLE 15-7 *Decomposition of racial/ethnic gaps in self-employment entry rates*

|  | Specification      |                       |                     |                   |
|--|--------------------|-----------------------|---------------------|-------------------|
|  | Native-born blacks | Native-born Hispanics | Hispanic immigrants | Native Americans  |
|  | (1)                | (2)                   | (3)                 | (4)               |
| Explanatory variables                                |                    |                       |                     |                   |
| White/minority gap in entry rate                     | 0.0144             | 0.0047                | 0.0071              | 0.0079            |
| Contributions from racial differences in:            |                    |                       |                     |                   |
| Sex  | -0.0002<br>-1.6%   | -0.0006<br>-16.8%     | -0.0009<br>-14.7%   | -0.0006<br>-17.2% |
| Education  | 0.0009<br>6.0%     | 0.0012<br>34.3%       | 0.0028<br>44.8%     | 0.0012<br>36.1%   |
| Age  | 0.0000<br>0.0%     | 0.0002<br>4.5%        | 0.0001<br>0.9%      | 0.0001<br>2.2%    |
| Marital status and children                          | 0.0007<br>5.0%     | -0.0002<br>-6.1%      | -0.0010<br>-16.3%   | -0.0009<br>-27.5% |
| Not employed   | -0.0005<br>-3.4%   | -0.0004<br>-10.9%     | -0.0008<br>-12.0%   | -0.0013<br>-37.1% |
| Assets   | 0.0022<br>15.5%    | 0.0018<br>55.4%       | 0.0028<br>45.0%     | 0.0019<br>54.2%   |
| Region   | 0.0010<br>6.7%     | -0.0018<br>-54.5%     | -0.0018<br>-28.1%   | -0.0027<br>-78.3% |
| Central city status                                  | 0.0008<br>5.4%     | 0.0010<br>29.4%       | 0.0013<br>20.5%     | -0.0023<br>-66.5% |
| Year effects   | 0.0001<br>0.6%     | 0.0001<br>2.9%        | 0.0000<br>0.8%      | 0.0002<br>5.7%    |
| All included variables (“explained” part of the gap) | 0.0049<br>34.0%    | 0.0013<br>27.1%       | 0.0025<br>35.6%     | -0.0043<br>-54.6% |

Notes: Current Population Survey, Matched Annual Demographic Surveys (1998–2003). (1) The sample consists of individuals (ages 25–55) who are not self-employed business owners in year *t*. (2) Contribution estimates are from non-linear decompositions. See text for more details.

The overrepresentation of blacks in regions of the country with low entry rates explains a modest portion of the gap. Also, the under-representation of blacks in rural areas, which have relatively high entry rates, contributes to the gap. Overall, racial differences in the explanatory variables explain roughly one third of the black/white gap in business creation rates. The remaining or “unexplained” portion of the racial gaps in self-employment entry rates may be due to lending discrimination and consumer discrimination against black-owned firms and/or the omission of important unmeasurable factors such as risk aversion.<sup>28</sup>

Table 15-7 also reports estimates for native-born and immigrant Latinos (reported in Specifications 2 and 3, respectively). The two most important factors in explaining the gaps between the two Latino groups and native-born

whites are assets and education. Relatively low levels of assets explain more than half of the entry rate gap for native-born Latinos and slightly less than half of the gap for immigrant Latinos. Apparently, low levels of assets are limiting opportunities for Latinos to start businesses and this factor, at least in percentage terms, is more important for Latinos than for blacks.

Relatively low levels of education among Latinos, especially immigrants, are also a limiting factor in business creation. A surprisingly high 53.1% of immigrant Latinos and 20.4% of native-born Latinos did not complete high school. Education differences account for 44.8% of the entry rate gap for Latino immigrants and 34.3% of the entry rate gap for Latino natives.

The under-representation of Latinos residing in rural areas also contributes to the gaps in entry rates. On the other hand, Latinos have a favorable regional distribution as evidenced by the negative contribution estimates. Latinos are disproportionately located in the West South Central, Mountain and Pacific regions where business entry rates are relatively high. This finding suggests that the entry rate gap would be even larger if Latinos had a similar geographical dispersion as whites. Similarly, entry rate gaps would be larger if not for the relatively high rates of nonemployment among Latinos and high entry rates into self-employment from nonemployment.

The entry rate into self-employment is 0.55% points lower among Native Americans than native-born whites. Low levels of education and assets are mainly responsible. Education and asset differences explain 36.1 and 54.1% of the gap, respectively. Although these factors alone explain nearly the entire gap in business creation rates, there exist a number of offsetting factors. The Native American regional composition, overrepresentation in rural areas, high levels of nonemployment, and family characteristics are favorable in terms of increasing business formation. These results imply that if Native Americans had similar geographical locations, family structures and levels of employment as whites the gap in entry rates would be substantially larger than that reported.

### *5.1. The Causes of High Exit Rates for Disadvantaged Minority Groups*

Table 15-8 reports estimates for the decomposition of exit rates from self-employment for native blacks, native Latinos, and immigrant Latinos. Sample sizes are relatively small for these groups (250–301 observations) because they condition on business ownership in the first survey year. Sample sizes for Native American business owners are too small to report estimates.

Native-born blacks are nearly twice as likely to leave self-employment annually as native-born whites. Although there is no clear dominant factor explaining the disparity in exit rates, racial differences in asset levels, region

TABLE 15-8 *Decomposition of racial/ethnic gaps in self-employment exit rates*

| Explanatory variables                                   | Specification      |                       |                     |
|---|--------------------|-----------------------|---------------------|
|   | Native-born blacks | Native-born Hispanics | Hispanic immigrants |
|   | (1)                | (2)                   | (3)                 |
| White/minority gap in exit rate                         | -0.1797            | -0.1135               | -0.1251             |
| Contributions from racial differences in:               |                    |                       |                     |
| Sex   | -0.0092<br>4.9%    | -0.0054<br>4.3%       | 0.0046<br>-4.0%     |
| Education   | -0.0059<br>3.2%    | -0.0074<br>6.8%       | -0.0267<br>20.7%    |
| Age   | -0.0066<br>3.6%    | -0.0086<br>7.9%       | -0.0057<br>4.4%     |
| Marital status and children                             | -0.0036<br>2.0%    | -0.0010<br>1.0%       | 0.0002<br>-0.2%     |
| Assets  | -0.0134<br>7.3%    | -0.0091<br>8.4%       | -0.0140<br>10.8%    |
| Region  | -0.0123<br>6.7%    | 0.0029<br>-2.7%       | 0.0029<br>-2.3%     |
| Central city status                                     | -0.0180<br>9.8%    | -0.0080<br>7.5%       | -0.0213<br>16.5%    |
| Year effects  | 0.0011<br>-0.6%    | -0.0013<br>1.1%       | -0.0004<br>0.3%     |
| All included variables<br>("explained" part of the gap) | -0.0679<br>37.8%   | -0.0379<br>33.4%      | -0.0604<br>48.2%    |

Notes: Current Population Survey, Matched Annual Demographic Surveys (1998–2003). (1) The sample consists of individuals (ages 25–55) who are self-employed business owners in year  $t$ . (2) Contribution estimates are from nonlinear decompositions. See text for more details.

distributions and central city status contribute to the gap. The overrepresentation of blacks in inner city areas, which have relatively high exit rates, provides the largest single contribution (9.8%). Racial differences in asset levels explain 7.3% of the gap, which is in the range of estimates from the PSID reported in Fairlie (1999). Estimates from the PSID indicate that 1.8 to 11.1% of the male black/white gap in exit rates from self-employment is explained by differences in asset levels. Recent estimates from the Characteristics of Business Owners (CBO) survey indicate that 43.2% of the gap in business closure rates is explained by differences in the amount of required startup capital (Fairlie and Robb, 2003), but the focus on businesses, startup capital and closure rates makes the results difficult to compare.<sup>29</sup>

Both native-born and immigrant Latinos have substantially higher exit rates than native-born whites. Lower levels of education and assets, a younger population of business owners and under-representation in rural areas partly explain why Latinos are more likely to leave self-employment. Education and

assets are especially important factors for Latino immigrants, explaining 20.7 and 16.5% of the gap in self-employment exit rates, respectively.

### *5.2. Other Potential Explanations for Ethnic/Racial Differences*

Additional factors that might explain low rates of entry and high rates of exit from self-employment among disadvantaged minority groups include, but are not limited to, racial differences in parental self-employment, sector-specific human capital and lending and consumer discrimination. Early researchers emphasized the role that past inexperience in business played in creating low rates of business ownership among blacks. In particular, Du Bois (1899), and later Myrdal (1944), Cayton and Drake (1946) and Frazier (1957) identified the lack of black traditions in business enterprise as a major cause of low levels of black business ownership at the time of their analyses. The lack of black traditions in business argument relies on a strong intergenerational link in business ownership. Indeed, several recent studies found that the probability of self-employment is substantially higher among the children of the self-employed (see Lentz and Laband, 1990; Fairlie, 1999; Dunn and Holtz-Eakin, 2000; Hout and Rosen, 2000).

Recent research has also examined whether the strong intergenerational link in business ownership is detrimental to disadvantaged minorities. Hout and Rosen (2000) noted a “triple disadvantage” faced by black men in terms of business ownership. They are less likely than white men to have self-employed fathers, to become self-employed if their fathers were not self-employed and to follow their father in self-employment. Fairlie (1999) provided evidence from the PSID that current racial patterns of self-employment are in part determined by racial patterns of self-employment in the previous generation. Finally, Fairlie and Robb (2003) found related evidence that the lack of prior work experience in a family business among black business owners, perhaps by limiting their acquisition of general and specific business human capital, negatively affects black business outcomes, such as closures, employment and sales. They also found that racial differences in business inheritance are negligible and cannot explain differences in outcomes.

Lending and consumer discrimination may also contribute to the patterns documented above. Recent evidence indicates that black-owned businesses experience higher loan denial probabilities and pay higher interest rates than white-owned businesses even after controlling for differences in credit-worthiness, size and other factors (Blanchflower, Levine and Zimmerman, 2003; Cavalluzzo, Cavalluzzo and Wolken, 2002). Minority-owned firms are also more likely to report not applying for loans because of concerns over being denied and that the availability of credit was a major problem. The evidence on consumer discrimination against minority-owned firms, however, is less clear (see, e.g., Borjas and Bronars, 1989; Meyer, 1990).

## 6. CONCLUSIONS

Estimates from the CPS indicate that several major disadvantaged groups have relatively low rates of entrepreneurship in the United States. Women, disadvantaged minorities (i.e., blacks, Latinos and Native Americans) and less-educated workers are found to have substantially lower business ownership rates than men, white non-Latinos and college-educated workers, respectively. An analysis of the dynamics of self-employment reveals some underlying causes of these patterns.

Although female self-employment rates have risen dramatically in recent decades, the prevalence of business ownership among women is only 50–60% of that for men. The low rate of self-employment among women permeates across ethnic/racial groups and countries. Evidence from U.S. and British data indicate that only a handful of detailed ethnic/racial groups have female/male self-employment rate ratios larger than 0.75 (Fairlie and Meyer, 1996; Clark and Drinkwater, 2000) and aggregate data from the OECD indicate that female self-employment rates are substantially lower than male rates in almost every reported country with an average ratio of 0.543.

Estimates from one-year transition matrices using matched CPS data indicate that women have lower rates of entry into self-employment from both nonemployment and wage/salary employment than men. Women also have substantially higher exit rates from self-employment than men. Slightly more than one-third of all self-employed women leave by the following year compared to one-fourth of self-employed men. These estimates imply that the low rate of business ownership among women is due to both a relatively low entry rate into self-employment and a relatively high exit rate out of self-employment.

Logit regressions for the probability of self-employment entry and exit are estimated to control for differences in ethnicity/race, immigration, education, nonemployment, assets, age, marital status, number of children, region of the country and central city status. As expected, the inclusion of these controls has little effect on the female/male entry and exit rate differentials because men and women generally have similar observable characteristics. Although previous research indicates that the determinants of self-employment differ between men and women (Bruce, 1999; Boden, 1996, 1999; Carr, 1996; Devine, 1994b; Lombard, 2001; Lohmann, 2001), the question of what explains the large gender gap in self-employment entry and exit remains largely unanswered. As reported above, evidence from the United States and several other countries suggests that women are less likely than men to report having a desire for self-employment, although the difference is not large (Kourilsky and Walstad, 1998; Blanchflower, Oswald and Stutzer, 2001). In the end, unobservable factors, such

as different preferences, discrimination and risk aversion, may be responsible for low levels of female entrepreneurship.

Estimates from the CPS indicate a clear ordering of self-employment propensities across ethnic and racial groups. White, non-Latinos and Asians have the highest self-employment rates followed distantly by Latinos. The likelihood of business ownership among Latinos is only slightly higher than 50% of that for white, non-Latinos. Native Americans have even lower levels of business ownership, and blacks have the lowest rates of business ownership, which are 36 to 39% of white rates. Aggregate data from the 2001 Canadian and U.K. Censuses and micro data from the 2000 U.S. Census indicate similar patterns of ethnic/racial entrepreneurship—disadvantaged groups, such as blacks, Latinos and Natives, have relatively low rates of business ownership in all of the countries reported. Thus, low rates of business ownership among these ethnic/racial groups are not peculiar to the United States or one country.

Blacks, Latinos and Native Americans are less likely to start businesses than are whites. All three groups are also more likely to leave self-employment. The differences in transition probabilities between these disadvantaged groups and whites are striking, especially for blacks. Only 1.2% of wage/salary blacks become entrepreneurs over a one-year period, which is less than half the white rate of entry into self-employment. For all three disadvantaged minority groups, exit rates are at least 40%, whereas the white exit rate is 26.8%. Clearly, low rates of business ownership among disadvantaged minorities are driven by both low entry rates and high exit rates.

To identify the contributions from ethnic and racial differences in education, assets and other factors to gaps in self-employment entry and exit rates a nonlinear decomposition technique is employed. For entry rates, the largest factor explaining disparities between native blacks and whites are racial differences in asset levels. Lower levels of assets among blacks account for 15.5% of the white/black gap in the probability of entry into self-employment. The two most important factors in explaining the gaps between native-born and immigrant Latinos and native-born whites are assets and education. Relatively low levels of assets explain more than half of the entry rate gap for native-born Latinos and slightly less than half of the gap for immigrant Latinos. Apparently, low levels of assets are limiting opportunities for Latinos to start businesses and this factor, at least in percentage terms, is more important for Latinos than for blacks. Relatively low levels of education among Latinos, especially immigrants, are also a limiting factor in business creation. Education differences account for 44.8% of the entry rate gap for Latino immigrants and 34.3% of the entry rate gap for Latino natives. Low levels of education and assets are also mainly responsible for the entry rate gap between Native Americans and native-born whites. Education differences explain 36.1% of the gap and asset differences explain 54.1% of the gap, however, there exist



many offsetting factors such as regional composition, overrepresentation in rural areas, high levels of nonemployment and family characteristics which are favorably associated with business creation.

The nonlinear decomposition technique is also used to identify factors explaining ethnic/racial differences in exit rates out of self-employment. Although there is no clear dominant factor explaining the black/white disparity in exit rates, racial differences in asset levels, regional distributions and central city status contribute to the gap. Lower levels of education and assets, a younger population of business owners and under-representation in rural areas partly explain why Latinos are more likely to leave self-employment. Education and assets are especially important factors for Latino immigrants, explaining 20.7 and 16.5% of the gap in self-employment exit rates, respectively.

A few recent studies have focused on an additional disadvantaged group—the less educated (see, e.g., Fairlie, 2004; Krashinsky, 2004). Estimates from the CPS indicate that only 6.5% of individuals who do not have a high school diploma are self-employed. In contrast, 11% of college-educated individuals own a business. The differences are smaller, however, after conditioning on employment. Estimates from transition matrices also reveal an U-shaped relationship between education and entry into self-employment from wage/salary employment. Entry into business ownership from nonemployment, however, is clearly increasing with education. Also contributing to the positive relationship between self-employment and education, exit rates from self-employment are found to be decreasing with education.

The addition of controls for ethnicity/race and assets in logit regressions is found to reduce the self-employment entry and exit rate differentials between the less educated and college graduates. The presence of liquidity constraints and relatively low levels of assets may limit the ability of less-educated workers to start businesses. Even controlling for differences in ethnicity/race and asset levels, however, the individual's education level has a large positive effect on entry into self-employment and a large negative effect on exit out of self-employment. High school dropouts are nearly a full percentage point less likely to enter self-employment and are 7.2 percentage points more likely to exit from self-employment than college graduates. Although many entrepreneurship programs targeted toward disadvantaged groups currently exist, the estimates presented here indicate continuing disparities in levels of business ownership.

## NOTES

<sup>1</sup> See OECD (1992) for descriptions of programs in Belgium, Canada, Finland, France, Greece, Netherlands, Portugal, Spain and the United Kingdom.

<sup>2</sup> See Bates (1993) for a description of programs promoting self-employment among minorities.

<sup>3</sup> See Glazer and Moynihan (1970), Light (1972, 1979), Sowell (1981) and Moore (1983).

<sup>4</sup> Job satisfaction is also much higher among the self-employed than wage/salary workers (Blanchflower, Oswald and Stutzer, 2001).

<sup>5</sup> A comparison of poverty rates reveals even more alarming differences. The Latino, black and Native American poverty rates range from 2.8 to 3.1 times the white, non-Latino poverty rate (U.S. Bureau of the Census, 2003).

<sup>6</sup> See Levy and Murname (1992) and Katz and Autor (1999) for reviews of the literature on wage inequality.

<sup>7</sup> Unpaid family workers are not counted as self-employed.

<sup>8</sup> Estimates from the CPS also indicate that women have substantially lower rates of self-employment than men even after stratifying the sample by the presence of children, marital status, full-time or part-time status, education level and age group.

<sup>9</sup> Estimates from the 1991 Census and the 1993–94 Fourth National Survey of Ethnic Minorities indicate similar ethnic and racial patterns in self-employment rates for Britain (Clark and Drinkwater, 1998, 2000).

<sup>10</sup> Estimates are similar excluding agricultural workers.

<sup>11</sup> Estimates from Canada indicate a slightly higher entry rate into self-employment from wage/salary work, a much higher entry rate into self-employment from nonemployment and a slightly higher exit rate out of self-employment for women than men (Kuhn and Schuetze, 2001). Their estimates, however, are not directly comparable because they count incorporated business owners as wage/salary workers.

<sup>12</sup> Krashinsky (2004) also notes high rates of entry into self-employment among less-educated workers displaced from their jobs.

<sup>13</sup> The main exceptions are that the high school dropout coefficient in the entry logit is no longer negative and statistically significant and the female and high school dropout coefficients in the exit logit are much smaller in magnitude.

<sup>14</sup> Estimates are similar if agricultural workers are excluded.

<sup>15</sup> The reported marginal effect provides an estimate of the effect of a one-unit increase in the independent variable on the self-employment entry probability. It equals the sample average of  $e^{X_i\hat{\beta}}/(1 + e^{X_i\hat{\beta}})$ .

<sup>16</sup> The exact causes of these differences are unknown and are beyond the scope of this chapter.

<sup>17</sup> For example, see Evans and Jovanovic (1989), Evans and Leighton (1989), Meyer (1990), Holtz-Eakin, Joulfaian, and Rosen (1994), Dunn and Holtz-Eakin (1999), Fairlie (1999, 2002) and Hurst and Lusardi (2004) for evidence from U.S. microdata, Holtz-Eakin and Rosen (2004) for U.S. and Germany, and Johansson (2000) for Finland.

<sup>18</sup> The focus on transitions to self-employment attempts to avoid the endogeneity problem of including assets in a static model of self-employment. A positive relationship found in a cross-sectional analysis may simply reflect the possibility that business owners accumulate more wealth instead of wealth increasing the likelihood of owning a business. Although individuals may save in anticipation of becoming self-employed, a measure of assets in the prior year should be more exogenous to the entrepreneurial decision than a contemporaneous measure of assets.

<sup>19</sup> Another approach that has been taken in the literature is to use inheritances, gifts, lottery winnings or insurance settlements as a measure of or instrument for assets (see Holtz-Eakin, Joulfaian, and Rosen, 1994a; Fairlie, 1999 and Hurst and Lusardi 2004 for U.S. microdata; Blanchflower and Oswald, 1998, and Taylor, 2001, for British microdata; and Lind and Ohlsson.

1994, for Swedish data). Inheritances and other unanticipated, or at least less-anticipated, lump sum payments represent a more exogenous measure of assets than net worth and are generally found to increase the probability of entering or being self-employed suggesting that entrepreneurs face liquidity constraints. Hurst and Lusardi (2004), however, found that future inheritances also increase the probability of self-employment entry suggesting that liquidity constraints are not the underlying cause of the positive relationship.

<sup>20</sup> Previous studies find that home prices, home ownership and property restitution increase the likelihood of business creation and self-employment (Fairlie, 2004; Black, de Meza and Jeffreys, 1996; Johansson, 2000; Earle and Sakova, 2000).

<sup>21</sup> Cross-sectional data for Europe indicates a negative relationship between education and self-employment (Blanchflower, 2004).

<sup>22</sup> Estimates from the National Center for Educational Statistics indicate that women received 49.6 and 40.7% of all Bachelor's and Master's degrees in business conferred in 2000–01 (U.S. Department of Education, 2002).

<sup>23</sup> See Gatewood et al. (2003) and Parker (2004) for recent reviews of the literature.

<sup>24</sup> Another possibility is that female entrepreneurs have access to different business and investment social networks than male entrepreneurs (Brush et al., 2004).

<sup>25</sup> See Coleman (2001) for a discussion of constraints faced by women-owned firms.

<sup>26</sup> The calculation of (3) and (4), however, is not possible without first matching the white distribution of  $X_1$  and the minority distribution of  $X_2$ . I drew a random subsample of whites with a sample size equal to  $N_M$  and match it to the minority sample based on the predicted probability of the dependent variable. To approximate the use of the entire white sample, I drew 1000 random white samples for matching and calculate the mean value of estimates from all of these matched samples. See Fairlie (2003) for more details.

<sup>27</sup> See Menchik and Jianakoplos (1997), Altonji and Doraszelski (2001) and Gittleman and Wolff (2004) for a few recent studies on racial differences in asset levels, and Bradford (2003) on wealth holding among black and white entrepreneurs.

<sup>28</sup> See Cavalluzzo, Cavalluzzo and Wolken (2002), Blanchflower, Levine and Zimmerman (2003), Borjas and Bronars (1989) and Meyer (1990) for evidence on lending and consumer discrimination against blacks, and see Fairlie (2002) for evidence on risk aversion.

<sup>29</sup> Using the 1982 CBO, Bates (1989) found that racial differences in levels of financial capital partly explain racial patterns in business failure rates.

## REFERENCES

- Altonji, J. G. and Doraszelski, U. (2001). *The Role of Permanent Income and Demographics in Black/White Differences in Wealth*. Working Paper No. 8473, National Bureau of Economic Research, Washington, DC.
- Aronson, R. L. (1991). *Self-Employment: A Labor Market Perspective*. ILR Press, Ithaca, NY.
- Aspen Institute (2002). *Directory of Microenterprise Programs*. The Aspen Institute, Washington, DC.
- Baron, S. W. and Kahan, A. (1975). *Economic History of the Jews*. Schocken Books, New York.
- Bates, T. (1989). The Changing Nature of Minority Business: A Comparative Analysis of Asian, Nonminority, and Black-Owned Businesses. *The Review of Black Political Economy* 18: 25–42.

- Bates, T. (1993). *Assessment of State and Local Government Minority Business Development Programs*. Report to the U.S. Department of Commerce Minority Business Development Agency, Washington, DC.
- Bates, T. (1997). *Race, Self-Employment & Upward Mobility: An Illusive American Dream*. John Hopkins University Press, Baltimore.
- Black, J., de Meza, D. and Jeffreys, D. (1996). House Prices, The Supply of Collateral and the Enterprise Economy. *Economic Journal* 106: 60–75.
- Blanchflower, D. G. (2000). Self-employment in OECD Countries. *Labor Economics* 7: 471–505.
- Blanchflower, D. G. (2004). Self-employment: More May Not Be Better. *Swedish Economic Policy Review* 11 (2): 15–74.
- Blanchflower, D. G., Levine, P. and Zimmerman, D. (2003). Discrimination in the Small Business Credit Market. *Review of Economics and Statistics* 85 (4): 930–943.
- Blanchflower, D. G. and Meyer, B. D. (1994). A Longitudinal Analysis of the Young Self-Employed in Australia and the United States. *Small Business Economics* 6 (1): 1–19.
- Blanchflower, D. G. and Oswald, A. J. (1998). What Makes an Entrepreneur? *Journal of Labor Economics* 16 (1): 26–60.
- Blanchflower, D. G., Oswald, A. and Stutzer, A. (2001). Latent Entrepreneurship across Nations. *European Economic Review* 45: 680–691.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Variables. *Journal of Human Resources* 8: 436–455.
- Boden, R. (1996). Gender and Self-Employment Selection: An Empirical Assessment. *The Journal of Socio-Economics* 25: 671–682.
- Boden, R. J. (1999). Flexible Working Hours, Family Responsibilities, and Female Self-employment: Gender Differences in Self-employment Selection. *The American Journal of Economic Sociology* 58 (1): 71–76.
- Bonacich, E. and Modell, J. (1980). *The Economic Basis of Ethnic Solidarity in the Japanese American Community*. University of California Press, Berkeley.
- Borooah, V. K. and Hart, M. (1999). Factors Affecting Self-Employment among Indian and Black Caribbean Men in Britain. *Small Business Economics*. 13: 111–129.
- Borjas, G. (1986). The Self-Employment Experience of Immigrants. *Journal of Human Resources* 21 (Fall): 487–506.
- Borjas, G. and Bronars, S. (1989). Consumer Discrimination and Self-Employment. *Journal of Political Economy* 97: 581–605.
- Bradford, W. D. (2003). The Wealth Dynamics of Entrepreneurship for Black and White Families in the U.S. *Review of Income and Wealth* 49 (1): 89–116.
- Brown, Ch., Hamilton, J. and Medoff, J. (1990). *Employers Large and Small*, Harvard University Press, Cambridge.
- Brush, C., Carter, N., Gatewood, E., Greene, P. and Hart, M. (2004). *Gatekeepers of Venture Growth: A Diana Project Report on the Role and Participation of Women in the Venture Capital Industry*. Kauffman Foundation.
- Bruce, D. (1999). Do Husbands Matter? Married Women Entering Self-Employment. *Small Business Economics* 13: 317–329.
- Carr, D. (1996). Two Paths to Self-Employment? Women's and Men's Self-Employment in the United States, 1980. *Work and Occupations* 23: 26–53.
- Carrasco, R. (1999). Transitions to and from Self-Employment in Spain: An Empirical Analysis. *Oxford Bulletin of Economics and Statistics* 61 (3): 315–341.
- Cavalluzzo, K., Cavalluzzo, L. and Wolken, J. (2002). Competition, Small Business Financing, and Discrimination: Evidence from a New Survey. *Journal of Business* 25 (4).
- Cayton, H. R. and Drake, S. C. (1946). *Black Metropolis*. Jonathan Cape, London.

- Clark, K. and Drinkwater, S. (1998). Ethnicity and Self-Employment in Britain. *Oxford Bulletin of Economics and Statistics* 60: 383–407.
- Clark, K. and Drinkwater, S. (2000). Pushed out or Pulled in? Self-employment Among Ethnic Minorities in England and Wales. *Labour Economics* 7: 603–628.
- Coleman, S. (2001). Constraints Faced by Women Small Business Owners: Evidence from the Data. Paper presented at the United States Association for Small Business and Entrepreneurship Annual Meetings.
- Constant, A. and Zimmerman, K. F. (2004). Self-Employment Dynamics across the Business Cycle: Migrants versus Natives. IZA Working Paper.
- Devine, T. J. (1994). Changes in Wage-and-Salary Returns to Skill and the Recent Rise in Female Self-Employment. *American Economic Review* 84 (2): 108–113.
- Devine, T. J. (1994). Characteristics of self-employed women in the United States. *Monthly Labor Review* 117 (3): 20–34.
- Du Bois, W. E. B. (1899). *The Philadelphia Negro*. University of Pennsylvania, Philadelphia.
- Dunn, T. A. and Holtz-Eakin, D. J. (2000). Financial capital, human capital, and the transition to self-employment: evidence from intergenerational links. *Journal of Labor Economics* 18 (2): 282–305.
- Earle, J. S. and Sakova, Z. (2000). Business start-ups or disguised unemployment? Evidence on the character of self-employment from transition economies. *Labour Economics* 7 (5): 575–601.
- Evans, D. and Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints. *Journal of Political Economy* 97 (4): 808–827.
- Evans, D. and Leighton, L. (1989). Some Empirical Aspects of Entrepreneurship. *American Economic Review* 79: 519–535.
- Fairlie, R. W. (1999). The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment. *Journal of Labor Economics* 17 (1): 80–108.
- Fairlie, R. W. (2002). Drug Dealing and Legitimate Self-Employment. *Journal of Labor Economics* 20 (3): 538–567.
- Fairlie, R. W. (2003). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. *Journal of Economic and Social Measurement* 30 (4): 305–316.
- Fairlie, R. W. (2004). Earnings Growth among Less-Educated Business Owners. *Industrial Relations* 43 (3): 634–659.
- Fairlie, R. W. (2004). Recent Trends in Ethnic and Racial Business Ownership. *Small Business Economics* 23: 203–218.
- Fairlie, R. W. (2004). *The Personal Computer and Entrepreneurship*. Working Paper, University of California, Santa Cruz.
- Fairlie, R. W. and Meyer, B. D. (1996). Ethnic and Racial Self-Employment Differences and Possible Explanations. *Journal of Human Resources* 31: 757–793.
- Fairlie, R. W. and Meyer, B. D. (2000). Trends in Self-Employment among Black and White Men: 1910–1990. *Journal of Human Resources* 35 (4): 643–669.
- Fairlie, R. W. and Robb, A. (2003). *Why are Black-Owned Businesses Less Successful than White-Owned Businesses: The Role of Families, Inheritances, and Business Human Capital*. Working Paper 336, Joint Center for Poverty Research.
- Frazier, E. F. (1957). *The Negro in the United States*, 2nd edition. McMillan, New York.
- Gatewood, E. J., Carter, N. M., Brush, C. G., Greene, P. G. and Hart, M. M. (2003). *Women Entrepreneurs, Their Ventures, and the Venture Capital Industry: An Annotated Bibliography*. Entrepreneurship and Small Business Research Institute, Stockholm.
- Gittleman, M. and Wolff, E. N. (2004). Racial Differences in Patterns of Wealth Accumulation. *Journal of Human Resources* 34 (1): 193–227.

- Glazer, N. and Moynihan, D. P. (1970). *Beyond the Melting Pot: the Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City*, 2nd edition. MIT Press, Cambridge, MA.
- Guy, C., Doolittle, F. and Fink, B. (1991). *Self-Employment for Welfare Recipients: Implementation of the SEID Program*. Manpower Demonstration Research Corporation, New York.
- Holtz-Eakin, D. and Rosen, H. (2004). *Cash Constraints and Business Start-Ups: Deutschmarks versus Dollars*. Working Paper, Syracuse University.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. (1994). Entrepreneurial Decisions and Liquidity Constraints. *Rand Journal of Economics* 23: 334–347.
- Hout, M. and Rosen, H. S. (2000). Self-Employment, Family Background, and Race. *Journal of Human Resources* 35 (4): 670–692.
- Hundley, G. (2000). Male/Female Earnings Differences in Self-Employment: The Effects of Marriage, Children, and the Household Division of Labor. *Industrial and Labor Relations Review* 54 (1): 95–104.
- Hundley, G. (2001). Why Women Earn Less than Men in Self-Employment. *Journal of Labor Research* 22 (4): 817–829.
- Hurst, E. and Lusardi, A. (2004). Liquidity Constraints, Household Wealth, and Entrepreneurship. *Journal of Political Economy* 112 (2): 319–347.
- Johansson, E. (2000). Self-employment and Liquidity Constraints: Evidence from Finland. *Scandinavian Journal of Economics* 102 (1): 123–134.
- Katz, L. F. and Autor, D. H. (1999). Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*, Volume 3, O. Ashenfelter and D. Card (Eds). North-Holland, Amsterdam.
- Kosanovich, W. T., Fleck, H., Yost, B., Armon, W. and Siliezar, S. (2001). *Comprehensive Assessment of Self-Employment Assistance Programs*. U.S. Department of Labor Report.
- Kourilsky, M. L. and Walstad, W. B. (1998). Entrepreneurship and Female Youth: Knowledge, Attitudes, Gender Differences, and Educational Practices. *Journal of Business Venturing* 13: 77–88.
- Krashinsky, H. (2004). *Self-Employment for Less-Educated Workers*. Working Paper, University of Toronto.
- Kuhn, P. J. and Schuetz, H. J. (2001). Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982–1998. *Canadian Economic Review* 34 (3): 760–784.
- Lentz, B. and Laband, D. (1990). Entrepreneurial Success and Occupational Inheritance among Proprietors. *Canadian Journal of Economics* 23: 563–579.
- Levy, F. and Murnane, R. J. (1992). U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations. *Journal of Economic Literature* 30: 1333–1381.
- Lindh T. and Ohlsson, H. (1996). Self-employment and Windfall Gains: Evidence from the Swedish Lottery. *Economic Journal* 106: 1515–1526.
- Light, I. (1972). *Ethnic Enterprise in America*. University of California Press, Berkeley, CA.
- Light, I. (1979). Disadvantaged Minorities in Self Employment. *International Journal of Comparative Sociology* 20 (1–2): 31–45.
- Light, I. and Rosenstein, C. (1995). *Race, Ethnicity, and Entrepreneurship in Urban America*. Aldine de Gruyter, New York.
- Lin, Z., Picot, G. and Compton, J. (2001). The Entry and Exit Dynamics of Self-Employment in Canada. *Small Business Economics* 15 (2): 105–125.
- Loewen, J. W. (1971). *The Mississippi Chinese: Between Black and White*. Harvard University Press, Cambridge.
- Lohmann, H. (2001). *Self-Employed or Employee, Full-Time or Part-Time? Gender Differences in the Determinants and Conditions for Self-Employment in Europe and the US*. Centre for European Social Research Working Paper 38, Mannheim, Germany.

- Lombard, K. V. (2001). Female Self-Employment and Demand for Flexible, Nonstandard Work Schedules. *Economic Inquiry* 39 (2): 214–237.
- Madrian, B. C. and Lefgren, L. J. (2000). An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents. *Journal of Economic and Social Measurement* 26: 31–62.
- Martinez-Granado, M. (2002). *Self-employment and Labor Market Transitions: A Multiple State Model*. Universidad Carlos III de Madrid Working Paper, Madrid, Spain.
- Menchik, P. L. and Jianakoplos, N. A. (1997). Black-White Wealth Inequality: Is Inheritance the Reason? *Economic Inquiry* 35 (2): 428–442.
- Meyer, B. (1990). *Why Are There So Few Black Entrepreneurs?* Working Paper No. 3537, National Bureau of Economic Research, Cambridge, MA.
- Min, P. G. (1989). *Some Positive Functions of Ethnic Business for an Immigrant Community: Koreans in Los Angeles*. Final Report Submitted to the National Science Foundation, Washington, DC.
- Min, P. G. (1993). Korean Immigrants in Los Angeles. In *Immigration and Entrepreneurship: Culture, Capital, and Ethnic Networks*, I. Light and P. Bhachu (Eds). Transaction Publishers, New Brunswick, Canada.
- Myrdal, G. (1944). *An American Dilemma*. Harper and Brothers, New York.
- Oaxaca, R. (1973). Male–Female Wage Differentials in Urban Labor Markets. *International Economic Review* 14: 693–709.
- Oaxaca, R. and Ransom, M. (1994). On Discrimination and the Decomposition of Wage Differentials. *Journal of Econometrics* 61: 5–21.
- OECD (1992). *Employment Outlook*. OECD, Paris. [Http://www.oecd.org/dataoecd/4/8/2409984.pdf](http://www.oecd.org/dataoecd/4/8/2409984.pdf).
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge University Press, Cambridge.
- Raheim, S. (1997). Problems and Prospects of Self-employment as an Economic Independence Option for Welfare Recipients. *Social Work* 42: 44–53.
- Razin, E. and Light, I. (1998). Ethnic Entrepreneurs in America's Largest Metropolitan Areas. *Urban Affairs Review* 33 (3): 332–360.
- Taylor, M. (2001). Self-Employment and Windfall Gains in Britain: Evidence from Panel Data. *Economica* 63: 539–565.
- U.S. Bureau of the Census (2003). *Poverty: 1999. Census 2000 Brief*. Washington, DC. [Http://www.census.gov/prod/2003pubs/c2kbr-19.pdf](http://www.census.gov/prod/2003pubs/c2kbr-19.pdf).
- U.S. Bureau of the Census (2004). *Evidence From Census 2000 About Earnings by Detailed Occupation for Men and Women, Census 2000 Special Reports*. Washington, DC. [Http://www.census.gov/prod/2004pubs/censr-15.pdf](http://www.census.gov/prod/2004pubs/censr-15.pdf).
- U.S. Department of Education (2002). *Digest of Education Statistics*. National Center for Education Statistics, Washington, DC.
- U.S. Small Business Administration (1998). *Women in Business*. Office of Advocacy, Washington, DC.
- U.S. Small Business Administration (1999). *Minorities in Business*. Office of Advocacy, Washington, DC.
- Vroman, W. (1997). *Self-Employment Assistance: Revised Report*. Urban Institute, Washington, DC.
- Walstad, W. B. and Kourilsky, M. L. (1998). Entrepreneurial Attitudes and Knowledge of Black Youth. *Entrepreneurship Theory & Practice* 23 (2): 5–18.

## Stage 6: Venture Performance and Harvesting



MARC COWLING

*Principal Economist, Institute for Employment Studies, Mantell Building, University of Sussex,  
Falmer, Brighton BN1 9RF, United Kingdom*

*E-mail: marc.cowling@ies.ac.uk*

## 16. Early Stage Survival and Growth

### 1. INTRODUCTION

It is widely argued that small businesses make a positive contribution to economies in two ways. First, they do so by enhancing productivity, primarily through introducing more competition into established markets, and by creating new markets through innovation and technological advance. Second, they do so by creating a disproportionate share of new employment opportunities. It thus follows that government should support small business, or at least consider the implications for small businesses, when bringing in new legislation, and indeed that they should enact policies favorable to the creation and sustainability of small business. As far back as Schumpeter (1942), a case was argued that small businesses are the vital mechanism by which women, ethnic minorities and immigrants enter economic activity via the labor market.

More recently, Robbins et al. (2000) expanded this line of argument by focusing on the role of small businesses in the secondary labor market which consists of first-time entrants into the formal labor market, the long-term unemployed, poorly educated individuals, young people, economically inactive women, minorities, immigrants, etc. They argue that big businesses deal in the primary labor market and small businesses have a much stronger presence in the secondary labor market. As such, when they create new employment opportunities, it is among relatively disenfranchised groups. In short, they can mop up difficult-to-shift groups of unemployed workers and provide greater opportunities for groups not typically in the mainstream, or formal, labor market. This creates, by implication, higher rates of economic activity as more people are drawn into the labor market, and lower levels of unemployment (Acs, 1999).

The evidence base for this is quite strong. For example, Bednarzik (2000), in his study of U.S. employment patterns over the business cycle in the early 1990s, found that employment patterns in established businesses were cyclical, yet new establishment formation remained unaffected over the full cycle. This supports the earlier case put forward by Binks and Jennings (1986) that small businesses insulate economies against recessionary effects by a combination of displaced large firm workers starting new businesses, and existing small businesses creating new employment opportunities. Bednarzik refers to this as the “smoothing effect.”

The employment effects are also backed by Swedish evidence from Folster (2000), who uses panel data for the period 1976–1995. Folster finds that self-employment raises regional employment, and that this effect is stronger with lags; Reynolds (1992) has shown that regional growth tends to be preceded by an increase in new business formation at the regional level.

From this, two important questions remain. Does a higher share of employment in the small business sector lead to productivity gains? And, is it true that small businesses create a disproportionately large share of net new employment? On the former, Robbins et al. (2000), in a U.S.-based panel study at the state level over the period 1986–1995, showed that the small business share of employment has a positive effect on productivity growth for businesses with fewer than 20 employees. This effect was not observed for larger SMEs. The employment share in very small businesses was also found to reduce wage inflation while, again, no effect was found for larger SMEs. They conclude that “very small businesses do indeed provide many of the benefits championed by small business proponents” (p. 300).

In a firm-level productivity study of 427 unquoted, independently owned British businesses, Cowling (2003) reported that there are very substantial productivity gains available to smaller firms by expanding their labor force, both at the managerial level and for normal workers. Yet, by implication, this also suggests that many smaller businesses are not as productive as they could be because they are too small. He also found that founding entrepreneurs do have a significant and positive effect on productivity. However, this effect is dissipated as firms grow large. This is consistent with decreasing returns to the fixed capabilities of the entrepreneur. Further, U.K.-based work examining productivity in 1,000 businesses found that small business (10–49 employees) is the most productive size class of firm (Harding et al., 2003).

Thus far we have outlined a case for promoting and supporting small business activity, and reported strands of evidence that the espoused benefits of new firm formation and small business activity do actually occur. For example, small businesses can increase productivity, and when they expand employment, it is likely to be in socially and economically beneficial ways. Yet small firms also have disadvantages. For example, they are associated with poorer working

conditions, less training and worse pay and promotion prospects. These issues need to be borne in mind when considering the apparent superior job creation capabilities of smaller firms.

So we now move on to consider whether small business activity results in more jobs at a more aggregate level. The body of evidence concerning whether or not small firms create jobs is fairly consistent. Whether they create more jobs than large firms has been the subject of intense scrutiny over a period of time, yet we cannot help but conclude that not only do small businesses create jobs, but that they create more jobs than large firms (see, e.g., Birch, 1979; OECD, 1996). In a U.S. study, Dennis (1993) reported that small businesses account for two-thirds of net new jobs. Despite a body of dissenting evidence (Davis, Haltiwanger and Shuh, 1996), and a technical debate around regression to the mean,<sup>1</sup> there is now a remarkable degree of consistency in the literature concerning the ability of newly established businesses to create jobs. In fact, Davidsson et al. (1998), in their job generation study prompted by the dissenting work of Davis et al. (1996), go so far as to argue that job creation policies should focus on newness whatever the size category of firm.

A related body of literature deals with Gibrat's Law and, in some cases, an alternative life-cycle model of firm growth. These approaches essentially explore the relationships between size and growth and, in the latter case, between size, age and growth. Broadly speaking, the earliest empirical studies, exclusively on large firm samples, showed that growth was independent of size or, in some cases, that the largest firms grew faster (see, e.g., Samuels, 1965; Prais, 1976). Yet for start-ups explicitly, Reid (1995), for new Scottish small firms in the 1980s, and Cowling and Williams (1998), for small U.K. firms in the mid 1990s, all found empirical evidence rejecting Gibrat's Law in favor of life-cycle models of firm growth. Evans (1987), for U.S. manufacturing, found that both hold. In short, younger firms grow faster. Heshmati (2000), using Swedish data for smaller firms with fewer than 100 employees, also found that age and size impact negatively on employment growth.<sup>2</sup>

Cabral (1995) associates this negative relation identified in the more recent studies, particularly those using samples of smaller firms, to the concept of sunk costs. Here, smaller and new firms that have positive exit probabilities find it optimal to underinvest in the initial start-up phase then, having survived, adjust their size to reach the long-run optimal level. This effect is observed as supranormal growth for smaller, surviving firms.

Having established that small firms as a group make a positive contribution to job creation and, in some cases, productivity, the key question that remains is what factors distinguish between those smaller businesses that experience growth and those that do not. Initially in this chapter, however, we focus on business survival. This is of great importance for several reasons. First, survival is the most basic measure of success, particularly in studies evaluating

the relative merits of business start-up support programs, and especially those dealing with unemployed start-ups. Second, it sheds light on the issue of whether the “wrong” kinds of people are starting up and continuing new businesses. This is valuable if our ultimate goal is to secure the highest growth from the start-up stock. Is it always the case that the “fittest” survive? Third, is it the case that governments are encouraging people who are likely to fail into making a decision that may have deleterious economic and social effects on them for many years into the future? Fourth, analyzing the particular characteristics of surviving and nonsurviving businesses may provide some important insights into areas where public policy might intervene to correct for market failures or imperfections. This in turn may lead to higher initial and subsequent growth. There is also an important technical issue surrounding survivor bias in many growth studies. Here, if we draw on survey evidence from existing businesses, we might arrive at conclusions that overestimate the actual growth of a particular cohort of businesses that started at a common point in time. This can occur if we fail to take into account the (negative) contribution to growth from nonsurviving businesses.

With these questions in mind, Section 2 examines the empirical evidence concerning the determinants of survival. It first discusses the impact of unemployment assistance programs on business survival before going on to explore what empirical studies tell us about the relative impacts of human capital, personal characteristics, business characteristics and macroeconomic conditions. Section 3 considers in depth the evidence about growth in new and early stage ventures. We explicitly focus on employment growth given its centrality to the political, social and economic case for supporting small firms, although we also consider other measures. In line with the basic structure of Section 2, we consider how human capital, personal and business characteristics, and macroeconomic conditions affect growth. But we will also explore how entrepreneurs’ competencies and strategic decision-making impacts on growth, as well as market based opportunities.

## 2. SURVIVAL: THE EVIDENCE BASE

It is clear that a combination of factors can influence the ability of new and small firms to create jobs. Useful summaries of potential factors are reported in Storey (1994) and Westhead and Cowling (1995). The models of Churchill and Lewis (1983), Scott and Bruce (1987) and Burns (1996) also provide some interesting insights into the early stages in the life of a new business (see Chapter 13 in this volume). All these models point to planning informality and the need to establish a customer base as characteristic of new businesses in the survival phase. They also point to the need to acquire

financial and business resources. In later (growth) phases strategic management skills are emphasized, as is the need for formalized structures and personnel skills to manage an expanded workforce. There is, broadly speaking, a shift from entrepreneurially managed business to formally managed business with functional disciplines such as marketing, accounting and finance, with a strong element of delegation. However, despite the intuitive appeal of such “stage” models, the actual sequence of issues outlined may not correspond to the empirical evidence. This will be investigated in detail subsequently.

A further issue of interest is that the weight of empirical evidence points to the fact that in a given sample of new businesses, which in total create large numbers of jobs, job creation is accounted for the most part by substantial growth rates in a very small subset of firms. These are often referred to in the literature as “gazelles.” The conclusion of this latter body of research is that gazelles have a combination of luck and the foresight to locate themselves in buoyant, often niche, markets which they then exploit by focusing on customers and quality. This points at an inherent difficulty in “picking winners” as the element of luck plays such a big role in determining future fast growth. But before we consider in detail various elements of early stage growth, we explore the business survival literature, as future growth is contingent upon surviving in business.

### *2.1. Business Survival*

Initially, we focus on existing research evidence on the impact of self-employment/business start-up programs. This is important as there has been a remarkable degree of support across countries for active labor market policies in this area. It also links back to our earlier discussion about the role of smaller businesses in secondary labor markets as many support programs of this nature focus very explicitly on promoting business start-up by the unemployed.

### *2.2. Unemployment and Public Assistance Programs*

The question of whether public money should be used to promote and support the unemployed to start up their own businesses is not without controversy. The key question is whether public money could be better used for alternatives that would provide a more efficient labor market outcome. For example, some previous research argues that if unemployed people who enter self-employment do so in response to poor labor market conditions that prevent them from obtaining waged employment, they may also lack the characteristics (such as financial and relevant human capital, and possibly attitudinal characteristics) necessary for starting, surviving and growing a new business.

It has also become more important for policy makers that the success or otherwise of assistance programs is seen in a wider labor market context. Thus, rather than measuring success simply as survival in self-employment, success is now seen as getting an individual into work, be it in self-employment or via self-employment into waged employment. This takes into account the fact that not all individuals would naturally choose self-employment, but if forced to enter through lack of jobs opportunities in the waged sector, the period of (self) employment may increase their future waged employment opportunities. In this context, Bruce and Schuetze (2004) found that there are no identifiable scarring effects of self-employment on future paid employment. Bryson and White (1996) provided an excellent review of earlier studies in this area, many of which reported lower survival rates for those entering self-employment from unemployment. Yet this early evidence contrasts with more recent, and often more rigorous, evidence (see, e.g., Cressy and Storey, 1995; Taylor, 1999).

Using cohort data on the unemployed, Bryson and White (1996) found that self-employed jobs for those previously long-term unemployed are more stable than comparable jobs in waged employment, although they also show that this effect dissipates over longer periods of time. Taylor (1999), using U.K. data from the British Household Panel Survey (BHPS), found that those entering self-employment since 1991 have a higher “survival” rate than those entering waged employment, and further, that this held for males and females. Another interesting finding from this work is that early exits from self-employment are nearly three times more likely to be voluntary shifts into better-paid employment. Cowling and Hayward (2000), analyzing data from 2700 individuals in a single U.K. locality who entered a business start-up program between 1991 and 1998, report similar evidence. Their findings show that of the 77.8% of program participants who actually started a new business and then subsequently shut it down, 43% entered waged employment. This result is comparable to that reported in Cowling and Taylor (2001), using BHPS data, who found that 44% of exits from self-employment found a waged job. It could be argued that the former evidence is stronger, given that the composition of the sample was comprised entirely of previously unemployed people. Further, in a follow-up study covering three U.K. localities, Cowling (2003) also shows that exiting from self-employment does not necessarily imply returning to, or entering, unemployment.

Taken together, the later evidence is consistent with the notion that experience of self-employment, and business start-up, may have an impact on an individuals’ labor market opportunities which is broadly similar to the effects of waged employment experience. There is also a further potential effect, which we explore subsequently, in that prior experience of self-employment and business start-up may improve an individuals’ chances of future success should they start a new business in the future. This is important as there is evidence that

increasing numbers of people active in the labor market have had multiple self-employment experiences. Cowling and Taylor suggested this may be as high as 20% compared to around 10% actively in self-employment at any given time. It also ties in with the literature relating to serial entrepreneurs (Westhead et al., 2003), who are people that continually set up new businesses.

Yet studies explicitly examining the (typically short-term) impacts of business start-up schemes for the unemployed tend to generate mixed findings. The literature (see Meager, 1996, for a good summary) has focused on the extent to which unemployed people, who were subsidized to start their own business, have survived in business. A subset of this body of literature also considered indirect job creation effects (i.e., how many jobs the subsidized self-employed have created for others); deadweight effects (i.e., whether scheme participants would have entered self-employment anyway); and displacement effects (i.e., whether new subsidized businesses have simply displaced activity from other, existing small businesses).

Broadly, this body of research finds that deadweight is often as high as 60–70%. Schemes also tend to attract the more advantaged unemployed (typically shorter-term unemployed, relatively highly qualified males, etc.). Displacement effects can also be high as subsidized businesses typically enter highly competitive markets with low margins and easy entry. And financial capital is an issue, both in terms of amounts invested and the point at which capital is invested. Cowling and Hayward (2000), for example, found that the majority of start-ups by unemployed people begin with less than £500 (\$750) invested in the business. Thus undercapitalization may put these businesses at a relative disadvantage and/or force them to set up in sectors with low capital requirements, low barriers to entry and a greater level of competition which may reduce survival and growth prospects.

Broadening the scope of our discussion beyond those studies that focus specifically on business start-up by the unemployed, or those which evaluate start-up programs by the unemployed, there is a large number of studies that consider factors which influence the likelihood of business survival and early stage business growth.

### *2.3. Human Capital*

Table 16-1 summarizes the findings of 19 studies conducted across five countries focusing on the determinants of business survival. From this empirical evidence, we observe that the age of the individual typically has a positive effect on the probability of business survival, although in certain cases this effect is no-linear, with survival rates dropping off in later years. This effect is generally interpreted as being a proxy for accumulated informal human capital, which can

TABLE 16-1 Review of empirical studies of factors affecting business survival

|                                 | Cowling and Mitchell (2004) U.K. | Meager et al. (2004) U.K. | Taylor (1999) U.K. | Van Praag (1994) U.S. | Evans and Leighton (1989) U.S. | Cooper et al. (1993) U.S. | Pfeiffer and Reize (1998) Germany | Cowling and Hayward (2000) U.K. | Cowling (2003) U.K. |
|---------------------------------|----------------------------------|---------------------------|--------------------|-----------------------|--------------------------------|---------------------------|-----------------------------------|---------------------------------|---------------------|
| <i>Human capital</i>            |                                  |                           |                    |                       |                                |                           |                                   |                                 |                     |
| Age                             |                                  | X                         | +                  | +                     |                                |                           | + then -                          | +                               | +                   |
| Education                       |                                  | +                         | X                  | X                     |                                | +                         |                                   |                                 | -                   |
| Self-employment experience      |                                  | +                         | +                  | X                     | +                              |                           |                                   |                                 |                     |
| Labor market experience         |                                  |                           | +                  | +                     |                                |                           |                                   |                                 |                     |
| Industry experience             |                                  |                           |                    | +                     |                                | +                         |                                   |                                 |                     |
| Unemployment experience         |                                  | X                         | -                  |                       |                                | +                         | X                                 | -                               | -                   |
| Personal characteristics        |                                  |                           |                    |                       |                                |                           |                                   |                                 |                     |
| Male                            |                                  | X                         | X                  |                       |                                | X                         | X                                 | X                               |                     |
| White                           |                                  | X                         | X                  |                       |                                | +                         |                                   | +                               | -                   |
| Disabled                        |                                  | X                         | X                  |                       |                                |                           |                                   | X                               | -                   |
| Occupation                      |                                  |                           | Yes                |                       |                                |                           |                                   | X                               |                     |
| <i>Business characteristics</i> |                                  |                           |                    |                       |                                |                           |                                   |                                 |                     |
| Industry                        | Yes                              | Yes                       | Yes                | Yes                   |                                | Yes                       | Yes                               | Yes                             | Yes                 |
| Age                             | +                                | +                         |                    |                       |                                |                           |                                   |                                 | +                   |
| Size                            | +                                | +                         |                    |                       |                                |                           |                                   | X                               | -                   |
| Capital                         | +                                | X                         |                    |                       |                                | X                         |                                   |                                 |                     |
| <i>Macroeconomic conditions</i> |                                  |                           |                    |                       |                                |                           |                                   |                                 |                     |
| Unemployment rate               |                                  |                           | +                  | X                     |                                | X                         |                                   | -                               | -                   |



TABLE 16-1 (Continued)

|                                 | Bates (1998) U.S. | Bosma et al. (2004) Netherlands | Cressy (1996) U.K. | Bruderl et al. (1998) Germany | Lin et al. (2000) Canada | Bruderl et al. (1992) Germany | Pennings et al. (1998) Netherlands | Cooper et al. (1994) U.S. | Bates (1990) U.S. | Bates (1995) U.S. |
|---------------------------------|-------------------|---------------------------------|--------------------|-------------------------------|--------------------------|-------------------------------|------------------------------------|---------------------------|-------------------|-------------------|
| <i>Human capital</i>            |                   |                                 |                    |                               |                          |                               |                                    |                           |                   |                   |
| Age                             | X                 | X                               | + then -           |                               | -                        |                               |                                    |                           | + then -          | + then -          |
| Education                       | -                 | X                               | +                  | +                             | +                        | +                             |                                    | +                         | +                 | +                 |
| Self-employment experience      | X                 | +                               | +                  | +                             |                          | X                             |                                    |                           |                   |                   |
| Labor market experience         |                   | X                               | +                  | +                             |                          | +                             |                                    | +                         |                   |                   |
| Industry experience             |                   | +                               |                    | +                             |                          | +                             | +                                  |                           |                   |                   |
| Unemployment experience         |                   |                                 | X                  |                               |                          |                               |                                    |                           |                   |                   |
| <i>Personal characteristics</i> |                   |                                 |                    |                               |                          |                               |                                    |                           |                   |                   |
| Male                            | +                 | +                               |                    |                               | +                        |                               |                                    | X                         |                   | X                 |
| White                           | X                 |                                 |                    |                               |                          |                               |                                    | +                         |                   | X                 |
| Disabled                        |                   |                                 |                    |                               |                          |                               |                                    |                           |                   |                   |
| Occupation                      |                   |                                 |                    |                               |                          | X                             |                                    |                           |                   |                   |
| <i>Business characteristics</i> |                   |                                 |                    |                               |                          |                               |                                    |                           |                   |                   |
| Industry                        | Yes               |                                 | Yes                | Yes                           |                          | Yes                           |                                    | No                        |                   |                   |
| Age                             | +                 |                                 |                    |                               | +                        | +                             | +                                  |                           |                   | -                 |
| Size                            | +                 |                                 | +                  | +                             |                          | +                             | +                                  |                           |                   | +                 |
| Capital                         | +                 |                                 | X                  | +                             |                          | +                             | +                                  | +                         |                   | +                 |
| <i>Macroeconomic conditions</i> |                   |                                 |                    |                               |                          |                               |                                    |                           |                   |                   |
| Unemployment rate               |                   |                                 |                    |                               | -                        |                               |                                    |                           |                   |                   |

Note: A '+' indicates variable was positive and significant (i.e., associated with higher survival). A '-' indicates a negative and significant relationship and 'X' that the variable has no statistically significant relationship.

be subject to decay in later years. However, there may also be a retirement effect which is not generally captured.<sup>3</sup>

Other measures of informal human capital often tested for are self-employment experience, more general labor market experience and industry-specific experience. On prior self-employment experience the evidence is fairly conclusive in that studies covering the U.K., U.S. and Netherlands all find a positive effect on survival. This strongly supports the notion that entrepreneurial human capital, possibly a learning-by-doing effect, raises the quality of the entrepreneurs input. More general work experience was also found to increase survival probabilities in a number of studies, as did industry specific experience. Taken as a whole, the empirical evidence strongly suggests that the most relevant form of human capital to individuals wishing to create a sustainable new venture is informal rather than formal. In short, it is experience accumulated through working (particularly in self-employment) that has the greatest impact on the ability of the entrepreneur to survive in business. The one caveat we add here is that in high-technology businesses, formal education plays a greater role as technology-based science requires specialist knowledge. Turning to the evidence relating to formal human capital, typically captured by years of completed schooling or highest completed educational qualification, the effects are mixed. Furthermore, this holds even within countries across different studies.

Returning to the issue of unemployment and business survival, we note that there is strong evidence that those who move into business from unemployment are significantly less likely to survive than those moving from waged employment. Further, given that one is unemployed, the longer that spell of unemployment was, the lower the probability of subsequent survival in business on one's own account. This is consistent with a decaying of skills and informal human capital the longer an individual is out of work. Cowling and Hayward (2000) found that the newly unemployed had a 20% higher survival probability than the very long-term unemployed. Yet, as we outlined previously, just having been in business, however successful, may enhance an individual's future employment prospects and, should they ever start another business, enhance the future performance of that business.

#### *2.4. Personal Characteristics*

Here we refer to the empirical evidence concerning the impact of personal characteristics on business survival. The most common demographics tested for are gender, ethnicity, disability and occupation (we discussed age in the context of human capital). Taken in order, we note that there is relatively little evidence of a gender impact on business survival, with the notable exceptions of the Dutch and Canadian studies by Bosma et al. (2004) and

Lin et al. (2000), respectively. This general result is interesting as it suggests that women are not underperforming on the most basic measure of business performance, that of survival. This implicitly leads us to question why there is such a large difference in male and female business start-up rates in most countries. But that line of inquiry is not within our remit here.

Ethnicity is another area in which we might expect to observe differences in business survival rates for a variety of factors. An important paper by Borjas and Bronars (1989) outlined how consumer discrimination might impact on self-employment outcomes. However, most of the literature relating to gender and ethnicity focuses on how women and ethnic minorities might have problems accessing the resources to start-up a business. Table 16-1 shows that most studies do not explicitly test for differences in survival rates across different ethnic groups. Of those that do, the evidence is mixed. For example, Taylor (1999) for the U.K. reports no differences. This contrasts with the two U.K.-based, studies reported by Cowling and Hayward (2000) and Cowling (2003) which, respectively, found positive and negative effects for whites. We can only conclude that there is a research gap and what little serious evidence exists is inconclusive on the issue of ethnicity.

Much the same can be concluded about any potential impacts on business survival from those with a disability. Very few studies explored this line of investigation, and of those that did most reported no impact, although one found a negative effect. This is possibly another area that warrants further investigation, particularly in light of increasing policy focus on economic and social inclusion, and the potential role that business support programs might play in promoting entrepreneurship among excluded groups.

Occupational status is another area that has received relatively sparse attention in the empirical literature. While there are potential problems that entrepreneurs sometimes classify themselves as owners when responding to a standard occupational classification question, rather than indicating that they are skilled manual (a construction business) or professional (lawyer), this is an area in which we might expect to observe significant variation and to possibly yield insights into transferability of skills. What little empirical evidence there is reports inconclusive results.

## *2.5. Business Characteristics*

Here we discuss findings from studies that consider basic business characteristics such as industry sector, age and size. These are all factors that we might expect to have a significant impact on business survival. For example, we might expect industry sector to capture differences in the level and structure of competition, barriers to entry and growth, scale economies and a host of other influences (see Chapter 7 of this volume). We know that smaller businesses tend

to have higher entry rates in construction and service sectors. And we also know that some of the most important new economy sectors, particularly knowledge based, are located in services. Thus even within the broad service sector we would expect to observe significant differences in survival rates.

The evidence is in line with our expectations as regards variations in survival rates across different industry sectors. In nearly all of the studies reported in Table 16-1, we observe statistically significant variation across industry sectors. Importantly, this holds across countries. Perhaps the most notable and consistent industry effect is for business services, which is generally found to be highly and positively associated with survival probability. Arising from this, an important issue is whether industry sector is a choice variable for individuals starting a business. In short, is it plausible that any person can start a new business in any sector they choose? This is a particularly important question in light of the evidence regarding industry-specific experience reported earlier which generally shows a positive impact on survival for those starting a business in an industry sector in which they have built up specific human capital through experiential learning.

While the majority of business survival studies do not include an age of business variable as they track businesses from inception, a small number of studies do. Of those that do, the results are consistent in that they all show a positive and significant effect indicating that the older a business is, the greater its survival probabilities. This is in line with a volume of evidence showing that the peak failure time is between 18 months and two years, a period during which approximately 80% of all failures occur. After this period, failure probabilities typically decline dramatically and after six years failure is not really an issue for the vast majority of surviving businesses.

Business size at start-up is also an important variable included in a number of empirical studies. *A priori*, we might predict that size is an indicator of resource availability, both financial and human capital, and of the quality of the entrepreneur or entrepreneurial team. As such, size should be associated with higher survival probabilities. Perhaps surprisingly, the empirical evidence is not conclusive on this issue. For example, Meager et al. (2003), Cowling (2003) and Cressy (1996), all for the U.K., and Bruderl et al. (1998) for Germany, report the expected positive relationship. But two other studies, one U.K. and one German, report no relationship, and a further U.K. study reports a negative relationship. The interesting feature is that all three of the latter studies are for previously unemployed people starting a new business. What this might suggest is that previously unemployed people might have more limited managerial skills, and that this may limit their capacity to manage larger-sized businesses. Thus, starting at a smaller scale may be more appropriate for certain types of entrepreneurs.

## 2.6. *Macroeconomic Conditions*

Finally, we consider the effect of the performance of the wider economy. The majority of studies that have tested for such effects have used the unemployment rate to proxy macroeconomic conditions. On balance this is probably a good indicator as it captures local demand for goods and services likely to be most relevant to new and early stage businesses. It might also be an important indicator of current and future inflows into self-employment, thus affecting the competitive environment that newer businesses are likely to face in their crucial, formative years. Turning to the empirical evidence, we note that the majority of studies that included this variable found a negative and significant effect from this variable. Thus, higher unemployment, particularly if this is captured at the local level, reduces business survival rates. This highlights an interesting dilemma in that government action in this area of the labor market is typically ratcheted up when unemployment is rising, or at historically high levels. As such we might question whether it is appropriate to encourage more business start-ups during periods when macroeconomic conditions are relatively unfavorable and survival chances are significantly lower.

## 2.7. *Summary of Empirical Evidence on Survival*

Having presented evidence from an array of empirical studies into factors associated with business survival from several developed countries, we now try to draw together the common themes. Perhaps the most consistent determinant of increased survival probabilities is the entrepreneur's age, although several studies report that this is nonlinear, implying that age acts in a positive way up to a certain point after which survival probabilities decline. Building on this human capital theme, there is further evidence that formal human capital and prior self-employment experience also enhances survival prospects. This is further enhanced if new entrepreneurs have relevant experience of the industry sector in which they set up their business. However, many of these positive human capital effects are dissipated if the individual spends a long time in unemployment prior to starting their business. To this end, policy might be usefully focused at developing human capital among that segment of the unemployed stock who might wish to start a new business in the future.

By contrast, the evidence is less conclusive regarding the impacts of gender, ethnicity, disability or occupational status. Indeed, if we broadly assume that an entrepreneurial society is one in which the small business population is similar in its demographic profile to the general working population, efforts might be better focused on promoting business start-up among groups that are

under-represented in the small business population via changing perceptions and culture.

We also note that there is a strong effect from the industry sector that an individual chooses to start their business in. In particular, “business services” appears to be a sector in which new and smaller businesses thrive. It is likely that higher survival rates are associated with being in sectors which are knowledge or human capital based rather than those in which financial capital, large volumes and economies of scale are important.

We also reaffirm the positive relationship between business age and size and survival. This former is consistent with many studies showing the inverted U-shaped relationship (with a long right-hand tail) between failure probabilities and business age, generally peaking around 18 to 24 months after start-up and declining strongly thereafter. Implicitly, we might assume that it takes this long for a bad entrepreneur to fully recognize his or her lack of entrepreneurial skills, or for initial reserves to be run down to such an extent that the business is no longer viable. This is in line with the Jovanovic’s (1982) learning-by-doing model in which entrepreneurial ability is revealed over time, with bad entrepreneurs exiting and good entrepreneurs surviving and growing as their abilities are revealed.

Importantly, given the structure of this chapter, which implicitly assumes a linkage between survival and growth, we also find that business size is commonly associated with survival. This is in line with the early findings of Phillips and Kirchoff (1989) who reported that new businesses that grow by even a small amount have dramatically higher survival rates. Thus, survival in business can be seen as an important indicator of performance and future growth potential, particularly if we consider that there is an expanding body of empirical evidence that shows persistence over time in terms of superior growth for initially high performing businesses (see, e.g., Cowling, 2004).

### 3. BUSINESS GROWTH: THE EVIDENCE BASE

This section reviews evidence about the determinants of early stage growth, that is, growth within the first five years after start-up. As in the previous section, we draw on studies that adopted a multivariate approach to examining the determinants of growth using relatively large samples of new and early stage businesses. Most of the studies cited refer to employment growth, although in some cases we draw on other work that has measured growth in other ways. In total, we consider more than 20 empirical studies covering a diverse set of countries including Israel, Canada, the U.S., U.K., Netherlands, Germany and France.

### 3.1. Human Capital

The first variable we consider is the age of the entrepreneur. Table 16-2 highlights the fact that, in the majority of studies that test for age effects, none are found. This contrasts with the survival results that showed a positive, and often nonlinear, effect. This suggests that general informal human capital, while enhancing survival probabilities, is not a good proxy for entrepreneurial talent as far as growing one's business in its formative years is concerned.

Thus, if we believe that entrepreneurial ability is a key factor in facilitating early stage growth, we must look elsewhere for variables that capture (or proxy for) these effects. One such indicator may be having a parent who is self-employed or running a business. This can be seen as capturing any informal entrepreneurial learning that may occur during one's early life, for example, working in one's parents business during school holidays or at weekends. The results are fairly inconclusive on this with relatively few studies incorporating this variable. But the Israeli study of small tourism ventures reported by Lerner and Haber (2000) found a positive effect, as does the U.K. work reported by Burke et al. (2000) using the National Child Development Study.

In a similar vein, prior self-employment experience is often cited as a key indicator of entrepreneurial ability, capturing all the learning effects of setting up and running a smaller business. Yet here again the results are anything but conclusive, although, perhaps surprisingly, this variable was not included in most studies. Of the four studies that did test for these effects, only two found a significant positive impact. These are the U.S.-based study of Westhead et al. (2003) and, once again, the U.K. National Child Development Study work of Burke et al. (2000).

While informal entrepreneurial human capital relevant to early stage growth is a very difficult concept to measure, there is limited evidence that industry experience (in the sector an individual sets their business up in) is associated with higher growth from two U.S. studies by Box et al. (1993) and Ensley et al. (1995). Yet there is virtually no evidence, in contrast to our survival findings, that prior experience of unemployment has a negative impact on early stage growth.

Next we consider the empirical evidence regarding the potential impact on early stage growth of formal education, our proxy for formal human capital. Here we observe that there is fairly strong empirical support, across a number of U.K.- and U.S.-based studies, for the notion that businesses with more educated entrepreneurs experience faster early stage growth. Furthermore, these studies also cover a reasonable time span, and different types of businesses, which might suggest that we can generalize with more confidence about this formal human capital effect.





TABLE 16-2 (Continued)

|   | Chandler et al. (1993) U.S. | Chandler et al. (1994) U.S. | Cooper et al. (1994) U.S. | Zahra (1996) U.S. | Robinson (1999) U.S. | Wesson (2001) U.S. | Lerner and Haber (2000) Israel | Durand and Coeurderoy (2001) France | Ensley et al. (2002) U.S. |
|---|-----------------------------|-----------------------------|---------------------------|-------------------|----------------------|--------------------|--------------------------------|-------------------------------------|---------------------------|
| <i>Human capital</i>                          |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Age   |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Parent self-employed                          |                             |                             | X                         |                   |                      |                    | +                              |                                     |                           |
| Education                                     |                             |                             | +                         |                   |                      |                    | X                              |                                     |                           |
| Self-employment experience                    |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Labor market experience                       |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Industry experience                           |                             |                             | X                         |                   |                      |                    |                                |                                     | +                         |
| Unemployment experience                       |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| <i>Personal characteristics</i>               |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Male  |                             |                             | +                         |                   |                      |                    |                                |                                     |                           |
| White   |                             |                             | +                         |                   |                      |                    |                                |                                     |                           |
| Disabled                                      |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Occupation                                    |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| <i>Business characteristics</i>               |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Industry                                      |                             |                             | Yes                       | X                 | X                    | +                  |                                | +                                   | +                         |
| Age   |                             | -                           |                           | X                 |                      |                    |                                | X                                   | X                         |
| Size  |                             |                             |                           | X                 |                      |                    |                                |                                     |                           |
| Capital                                       |                             |                             | +                         |                   |                      |                    |                                |                                     |                           |
| <i>Macroeconomic conditions</i>               |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Unemployment rate                             |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| <i>Competencies, strategy and opportunity</i> |                             |                             |                           |                   |                      |                    |                                |                                     |                           |
| Entrepreneurial competence                    | +                           |                             |                           |                   |                      |                    |                                |                                     |                           |
| Managerial competence                         | X                           |                             | X                         |                   | +                    |                    | +                              | +                                   | +                         |
| Strategy                                      |                             | +                           |                           | +                 | +                    |                    | X                              | +                                   | +                         |
| Opportunity                                   |                             | +                           |                           | +                 | +                    | +                  | +                              | +                                   |                           |

Note: A '+' indicates variable was positive and significant (i.e., associated with higher growth). A '-' indicates a negative and significant relationship and 'X' that the variable has no statistically significant relationship.

### 3.2. *Personal Characteristics*

Next we review the impact of personal characteristics on early stage growth. This is important as there has been a lot of debate about discrimination against women, ethnic minorities and, in some cases, those with a disability, and policy-makers have reacted to this perceived discrimination by implementing support programs to correct for these perceived disadvantages.

Perhaps the most interesting feature to note is that relatively few studies actually test for these effects. The U.S.-based study of Sapienza et al. (1997) reported no gender effect, while Cooper et al. (1994) found a positive effect for males. In European studies, Cowling (2002) in an EU-wide study found a positive effect for males, which is in line with the Bosma et al. (2004) Netherlands study and the Bruderl et al. (1998) German study. Only Cowling (2003) found a positive female effect for those using a publicly funded business start-up program in deprived areas of England.

Yet the interpretation of negative female effects is not clear-cut and does not necessarily imply female underperformance due to some innate male superiority. An interesting study in Sweden conducted by Du Rietz and Henrikson (2000), among 4200 independent businesses with between one and 20 employees, showed that female entrepreneurs are less likely to have real opportunities for expanding their businesses than males. Thus it is a dearth of growth opportunities that appears to be driving female entrepreneurs' lower growth outcomes. The critical piece of the jigsaw from Du Rietz and Henrikson's study was that if women entrepreneurs were faced with the same growth opportunities as their male counterparts then they would be equally likely to take them. However, the actual survey variable is a self-reported measure asking entrepreneurs "are there good prospects for expansion?" Thus it may be that female entrepreneurs are not identifying growth opportunities that actually exist. From this, we might conclude that this line of inquiry merits serious research consideration to fill this apparent gap in our understanding.

Concerning other personal characteristics, the empirical evidence is significantly less voluminous. On ethnicity, for example, only Cooper et al. (1994) for the U.S. and Cowling (2003) for deprived areas of England, find any ethnicity impacts. In both cases they identified a positive effect for white people. This contrasts with the U.K.-based study of young people starting a business of Meager et al. (2003) which found no such effect. On disability, the evidence is extremely limited, and those studies that do test for such effects generate mixed results. In the study by Cowling (2003), having a disability was found to exert a negative impact on early stage growth. Yet in the U.K. study of Meager et al. (2003), which looked at the experiences of young people, the reverse was found. Thus we cannot help but conclude that here again there is a significant gap in our knowledge and understanding about relative growth rates of ethnic minority

businesses compared to white-owned businesses, and also on the question of whether disability matters.

### 3.3. *Business Characteristics*

The first issue we consider is whether or not industry sector has an identifiable effect on early stage growth. As discussed previously in relation to survival, there are many reasons why we might expect to observe an empirical relationship of this sort. These include economies of scale, barriers to growth, competition and overall market growth. In line with our *a priori* thinking, we note that in a majority of studies that have tested for any such effects a significant industry effect is apparent. The most common sectors associated with higher growth rates are businesses services and manufacturing. And those associated with lower growth rates are personal household and other services. Reassuringly, this result holds across countries (see Durand and Coeurderoy, 2001, for French evidence; Cooper et al., 1994, for U.S. evidence; and Meager et al., 2003, for U.K. evidence).

Although our focus is on early stage growth, we might also expect the age of the business to affect realized growth. We know from the survival literature that most businesses struggle in their formative years, and many fail within two years of starting up. The empirical evidence on early stage growth generally supports this transition between survival mode and growth in that we observe a significant and positive effect for business age in a nearly half of all the growth studies that tested for these effects. However, some studies identified a negative effect (see Chandler et al., 1994a and 1994b, for U.S. evidence; and Durand and Coeurderoy, 2001, for French evidence). Other studies in the U.K. and U.S. failed to identify any such effect.

### 3.4. *Competence, Strategy and Opportunity*

In this section we consider how researchers have tried to capture aspects of entrepreneurial and managerial competence, strategic positioning, and the presence of business opportunity. Specifically, we explore whether these factors have been found to have any identifiable effect on early stage growth. As is common in the strategy literature, researchers have often adopted a bundling approach in order to identify complementary strategies and reduce the volume of data. The general procedure would be to use Cronbachs alpha to identify a strategy or competency index usually under the assumption that it is more likely that a bundle of complementary strategies, an indicator of systematic decision-making, will have an impact on performance than isolated, or non-joined up strategic decision-making. Other common procedures adopted are factor analysis and principal components.

First we consider entrepreneurial competence. This has been tested in for in a number of studies. Two U.S. studies, both by Chandler et al. (1993, 1994) found a positive correlation between entrepreneurial skills and competencies and business growth. By contrast, Sapienza et al. (1997) and Westhead and Cowling (1995) found no such association. Robson and Bennett (2000) found that informal external advice and being a member of a professional association has a positive impact on growth, suggesting that human capital and entrepreneurial competency can be developed and enhanced by interacting with others outside the business. Yet Harnes and Senneseth (2001) found that networking had no impact on employment growth in their analysis of 1700 businesses across eight European cities. Interestingly, Reid and Smith (2000) found that formal business planning had a negative impact on growth among Scottish businesses, although contrasting findings were reported by Van Gelderen et al. (2000).

Entrepreneurial goals were also found to have an impact in the Reid and Smith study, although with contrasting results. For example, running your business as a hobby was found to have a negative effect, as was a desire to be your own boss, and just seeking to survive in business. Yet having a profit goal was positively associated with business growth. The Meager et al. (2003) study of young people in business finds that entrepreneurial incomes are higher for those with an explicit growth objective. Interestingly, they also tested for the effects of risk-taking propensity and business mentoring on employment growth and found no significant relationships.

Focusing on studies that have examined potential impacts on early stage growth from managerial competencies we find a number of interesting results. For example, Chandler et al. (1994), Lerner and Haber (2000) and McGee et al. (1995) all found that greater competence at the managerial level is positively associated with business growth. The Scottish work of Reid and Smith (2000) also found that organizational capability enhances growth, and the case study work on e-commerce based SMEs of Feindt et al. (2002) linked having defined processes and product presentation to better performance. Once again, however, we found some studies that failed to identify such an association (Westhead and Cowling, 1995; Chandler et al., 1993) and some that identified a negative relationship (Sapienza et al., 1997).

Thus there is an interesting tension apparent between formality of business planning and general managerial competence. While we might assume that businesses at this stage of their life-cycle are entrepreneurially managed (i.e., informally), it is also the case that this needs to be aligned to more general managerial competence. This is particularly so when, having survived, growth becomes not only a desirable objective but a realistic opportunity. The tension lies in how entrepreneurial owned and managed businesses integrate new managers into the business and set up systems that can support managerial

processes and decision-making, but not stifle the informality, creativeness and speed of decision-making that characterizes early-stage entrepreneurial businesses and is arguably their greatest asset.

The impact of strategic decision-making has been addressed in a relatively large number of empirical studies across many countries, including the U.S., U.K., France, Germany, Netherlands and Israel. On balance, the results strongly indicate that strategy has an impact on growth. Thus the strategic decision-making processes of small, early-stage businesses are a critical factor in the determination of subsequent performance and growth, even in the face of strong, exogenous, market forces. On the assumption that the founding entrepreneur, or entrepreneurial team, are likely to determine strategic direction in businesses at this stage of their development, it is perhaps at this level that we are most likely to capture the essence of the entrepreneur and what they do. For it is the (superior) decision-making capabilities of entrepreneurs that fundamentally determine the shape, direction and outcomes of their businesses.

Empirically, there appears to be a reasonable body of evidence that suggests that reactive, or follower, strategies are associated with lower growth (Bruderl and Preisendorfer, 1998; Van Gelderen et al., 2000). Reid and Smith (2000) found that the extent of the strategic horizon (the time period over which strategic decisions are intended to cover) is critical to growth, and Feindt et al. (2002) show a positive association between managing customer relationships and success in E-commerce. Thus it does appear that the outward market-facing role of entrepreneurs and entrepreneurial decision-making is a vital element in the determination of early stage performance. However, not all of the studies that tested for strategic impact found it to be important. Here we note that those of Sapienza et al. (1997), Lerner and Haber (2000) and Westhead and Cowling (1995) all found little evidence that strategic positioning mattered.

Finally, we consider the body of empirical studies that measured opportunity, a necessary, if not sufficient, condition for growth. On this issue we have, bar one study, a strong consensus in terms of clearly identifying a positive association between the availability of opportunity and actual realized growth. Yet, as was apparent from the Swedish work of Du Rietz and Henriksen (2000), opportunity is not evenly distributed across the entrepreneurial population. Yet even in this study, which found that female entrepreneurs were less likely to have growth opportunities, when they did present themselves, female entrepreneurs were just as likely to take them up as male entrepreneurs.

### *3.5. Summary of Evidence on Growth*

Building on the empirical evidence concerning the ability of new businesses to survive, in this section we moved on to consider the determinants of growth. While we began by making the point that initial survival is a necessary

precondition for future growth, it is growth *per se* that is the *raison d'être* behind the huge political and popular support for entrepreneurs and entrepreneurship. So from the substantial body of evidence on early-stage growth, what themes emerge most consistently in the literature, and can we pull these together in a coherent way that will encapsulate what we know about high-performing, early-stage businesses?

On personal characteristics, the findings are inconclusive regarding gender effects, although there is marginally more evidence in support of female under-performance. Further, the age of entrepreneurs was not generally found to be an important factor. Where it was, the effect was nonlinear, with growth being positively associated with age up to a point after which growth declined as entrepreneurs approached the last decade before retirement. Formal education was also associated with higher growth, yet measures of informal human capital (prior self-employment experience, prior industry experience, etc.) tended not to have a decisive effect, perhaps surprisingly.

But a critical factor, and one about which there was a large degree of consistency across studies, was that industry sector matters. There was also evidence that for early-stage businesses, just being in business for a few more years can lead to higher growth rates. This is consistent with a shift from survival mode to growth orientated as it becomes less likely that the business will fail.

Finally, we note that measures of entrepreneurial and general managerial competency were not robustly associated with higher growth. Yet strategy and opportunity were. Thus it is likely that both entrepreneurial and management competencies are best captured through what the entrepreneur or entrepreneurial team do than trying to measure it a step back in a more formal way. On the evidence discussed here, we are drawn to the conclusion that it is the decision-making abilities of people that fundamentally determines who is a good entrepreneur and who is not. Although we might argue that better entrepreneurs are also more likely to spot opportunities, this entrepreneurial strategic input is most effective when growth opportunities exist in the marketplace.

To this end, we can suggest that a potentially rich strand of research might investigate in greater depth the experiences and competencies of entrepreneurs who make superior strategic decisions that then lead to growth. We have also identified other gaps in our knowledge. For example, there is a dearth of rigorous empirical work surrounding ethnicity and disability and growth. This may have led policy makers to make potentially ill-informed judgments about the need for support programs when swayed by politically expedient motives. What we can say is that if an entrepreneur (or team of entrepreneurs) selects the right industry sector to locate their business in, and then designs an appropriate strategy to ensure first of all that they survive for the initial few years, then reorientate themselves to be in a position to take advantage of growth

opportunities as they present themselves, then it is likely that these business will outperform the rest.

#### 4. CONCLUSION

We began by outlining the case for small businesses in terms of making a positive contribution to economies on two fronts. First, they enhance productivity and competition. Second, they enable relatively disadvantaged population groups to enter economic activity via the labor market. We then considered the evidence base and were drawn to the conclusion that small businesses do seem to provide many of the espoused benefits attributed to them.

On job creation, we observe a remarkable degree of consistency in the literature, albeit with a few dissensions, concerning the ability of newly established businesses to create net new jobs, although this may in part be explained by sub-optimality in size at the start-up point. Yet we also noted that most growth and jobs are actually accounted for by a very small subset of fast growth businesses, often termed gazelles. We also established that it was not appropriate to consider growth of early stage businesses without dealing with issues surrounding initial survival. Thus the framework of the chapter was such that we reviewed the survival literature, before going on to consider and review the early stage growth literature.

Regarding survival, we found that the most consistent factor across a diverse range of studies was the age of the entrepreneur, which tended to increase survival probabilities. However, we also noted that in many cases the effect was nonlinear and subject to decline in later years. Other factors that captured aspects of both formal and informal human capital were also found to be important, for example, prior self-employment experience and industry experience. Yet a lengthy spell of unemployment was found to substantially decrease survival probabilities. This suggests that human capital is critical to survival, but that it is also subject to decay if an individual is out of the active labor force for any length of time.

This suggests that retraining or skills updating might be a critical factor in improving the success of public policy support programs aimed at promoting new business start-ups from among the unemployed and those outside the active labor market. In short, it is current human capital that matters, not historical human capital accumulation. This is important given the huge level of public expenditure across the world on support programs of this nature, and the findings of many evaluations, which show relatively high failure rates, displacement effects and deadweight costs.

Yet we find less inconclusive evidence concerning the effects of gender or ethnicity, and are drawn to the conclusion that the imbalance we observe in

female entry rates might be largely explained by misperceptions and cultural differences rather than market based discrimination *per se*. It is also apparent that there is a large research gap in terms of our understanding how disability affects business entry, survival and performance. To a lesser extent, the same is true for ethnicity.

It was also apparent that there is huge variation in survival rates across different industry sectors. If we accept that industry sector is a choice variable for prospective entrepreneurs (and it is an important question whether it is), then individuals need to think carefully about their decision. Where there is a degree of consistency in the empirical literature is that sectors that are knowledge- (human capital-) based, rather than those in which economies of scale are important, have higher survival rates.

Turning to our review of the early stage growth literature, we sought to build upon our review of the business survival literature by considering whether the same factors that enhance initial survival also promote higher growth. This is important as growth is not evenly distributed among new businesses, and may require a distinct set of competencies in order to identify opportunities, set the appropriate strategies in place and manage the whole process. From the basic evidence, we know that the majority of start-ups begin and remain at a very small scale.

Regarding the impact of personal characteristics, we note that there is some evidence to support the female under-performance hypothesis in respect of growth. Importantly, this has been attributed to lack of opportunity, which was not the case for survival. Further, the age of the entrepreneur appears to be far less important for growth than survival. Here we observe quite an important and interesting shift in terms of defining the most important measure of human capital for survival and growth. For survival, the evidence suggests that informal human capital is the dominant measure. Yet for growth, formal educational measures appear to exert a greater influence. This is certainly an issue that merits further consideration in future work.

It is also very apparent that superior strategic decision-making is fundamental to growth. Thus we might conclude that the true measure of entrepreneurial human capital, in a growth context, is best captured through identifying strategic choice variables that distinguish between high and low growth businesses. This might also be enhanced by considering the structure of corporate governance in smaller, early stage businesses. This would allow us a better understanding of how top-level strategic decision making is conducted and operationalized in the business.

We conclude by stating that there are substantial and important differences in terms of the critical factors needed for new businesses to survive in the first instance and then embark on a growth trajectory. This can, at least partly, explain why comparatively more firms survive than achieve high and



sustainable growth. However, we also note that if entrepreneurs choose to enter the wrong industry to start with, then it becomes a virtual certainty that they will end up struggling to survive or grow. There is also a strong case to be made for entrepreneurs with low levels of formal human capital to bring educated managers and decision makers into the firm at an early stage.

## NOTES

<sup>1</sup> The basic issue here is one of when the initial measurement period occurs. If, for example, initial size measurement happened to be during an unnaturally low period in the firm's life, that is, below trend size, then when the firm adjusts its size back to its natural scale any growth rate identified would be an over-estimate of real underlying growth. One way to counteract this is to take the average of sizes measured across the period and calculate end period growth as the difference between that and the average size.

<sup>2</sup> There is also an issue about zero bias if firms who begin life with no additional employees then hire one more person. Then the growth rate is clearly large. To correct for this some studies have used actual employment change in numbers (see, e.g., the review of employment growth studies in Westhead and Cowling, 1995).

<sup>3</sup> I thank an anonymous referee for highlighting this important point.

## REFERENCES

- Acs, Z. (1999). *Are Small Firms Important? Their Role and Impact*. Kluwer Academic, Boston, MA.
- Bates, T. (1990). Entrepreneurs Human Capital Inputs and Small Business Longevity. *Review of Economics and Statistics* 72: 551–559.
- Bates, T. (1995). A Comparison of Franchise and Independent Small Business Survival Rates. *Small Business Economics* 7 (5): 377–388.
- Bates, T. (1998). Survival Patterns Among Newcomers to Franchising. *Journal of Business Venturing* 13 (2): 113–130.
- Bednarzik, R. (2000). The Role of Entrepreneurship in U.S. and European Job Growth. *Monthly Labor Review* 123 (7): 3–16.
- Binks, M. and Jennings, A. (1986). Small Firms as a Source of Economic Rejuvenation. In *The Survival of the Small Firm: The Economics of Survival and Entrepreneurship*, J. Curran, J. Stanworth and D. Watkins (Eds). Gower Publishing, Aldershot.
- Birch, D. (1979). *The Job Generation Process*. MIT Press, Cambridge, MA.
- Borjas, G. and Bronars, S. (1989). Consumer Discrimination and Self-Employment. *Journal of Political Economy* 97: 581–605.
- Bosma, N., van Praag, M., Thurik, R. and de Wit, G. (2004). The Value of Human and Social Capital Investments for the Performance of Start-Ups. *Small Business Economics* 23: 227–236.
- Box, T., White, M. and Baw, S. (1993). A Contingency Model of New Manufacturing Firm Performance. *Entrepreneurship, Theory & Practice* 18 (2): 31–46.
- Bruce, D. and Schuetze, H. (2004). The Labor Market Consequences of Experience in Self-Employment. *Labor Economics* 11 (5): 575–598.

- Bruderl, J. and Preisendorfer, P. (1998). Network Support and the Success of Newly Founded Businesses. *Small Business Economics* 10 (3): 213–225.
- Bruderl, J., Preisendorfer, P. and Ziegler, R. (1992). Survival Chances of Newly Founded Business Organizations. *American Sociological Review* 57: 227–242.
- Bryson, A. and White, M. (1996). *From Unemployment to Self-Employment*. Policy Studies Institute, London.
- Burke, A., Fitzroy, F. and Nolan, M. (2000). When Less is More: Distinguishing between Entrepreneurial Choice and Performance. *Oxford Bulletin of Economics and Statistics* 62 (5): 565–587.
- Burns, P. (1996). Growth. In *Small Business and Entrepreneurship*, P. Burns and J. Dewhurst (Eds). Palgrave, London.
- Cabral, L. (1995). Sunk Costs, Firm Size and Firm Growth. *Journal of Industrial Economics* 43: 161–172.
- Carroll, R., Holtz-Eakin, D., Rider, M. and Rosen, H. (2000). Income Taxes and Entrepreneurs Use of Labor. *Journal of Labor Economics* 18 (2): 324–351.
- Chandler, G. and Hanks, S. (1993). Measuring the Performance of Emerging Business. *Journal of Business Venturing* 8 (5): 391–408.
- Churchill, N. and Lewis, V. (1983). The Five Stages of Small Business Growth. *Harvard Business Review*: 30–40.
- Cooper, A., Gimeno-Gascon, F. and Woo, C. (1994). Initial Human and Financial Capital as Predictors of New Venture Performance. *Journal of Business Venturing* 9 (5): 371–395.
- Cowling, M. (2002). *The Contribution of the Self-Employed to Employment in the EU*. Small Business Service Research Report URN 03/539; London; [www.sbs.gov.uk/research/](http://www.sbs.gov.uk/research/).
- Cowling, M. (2003). Productivity and Corporate Governance in Smaller Firms. *Small Business Economics* 20: 335–344.
- Cowling, M. (2003). *Creating Local Opportunity*. Research Report No. 177, Department for Work and Pensions, London ([http://www.dwp.gov.uk/resourcecentre/archive/research\\_analysis\\_pubs.asp](http://www.dwp.gov.uk/resourcecentre/archive/research_analysis_pubs.asp)).
- Cowling, M. (2004). The Growth–Profit Nexus. *Small Business Economics* 22 (1): 1–9.
- Cowling, M. and Hayward, R. (2000). *Out of Unemployment*. Research Centre for Industrial Strategy, Birmingham University. ISBN: 0704421577.
- Cowling, M. and Mitchell, P. (2003). Is the Small Firms Loan Guarantee Scheme Hazardous for Banks? *Small Business Economics* 21: 63–71.
- Cowling, M. and Williams, C. (1998). *Growth: Size, Age and Productivity Effects*. Research Centre for Industrial Strategy, Birmingham Business School.
- Cressy, R. (1996). Are Business Start-Ups Debt Rationed? *Economic Journal* 106: 1253–1270.
- Cressy, R. and Storey, D. (1995). *New Firms and their Bank*. National Westminster Bank, London.
- Davidsson, P., Lindmark, L. and Olofsson, C. (1998). The Extent of Overestimation of Small Firm Job Creation: An Empirical Investigation of the Regression Bias. *Small Business Economics* 11: 87–100.
- Davis, S., Haltiwanger, J. and Shuh, S. (1996). Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts. *Small Business Economics* 8 (4): 297–315.
- Dennis, W. (1993). *Small Business Prime: Charts and Graphs Illustrating the Importance and Role of American Small Business*. NFIB Foundation, Washington, DC.
- Durand, R. and Coeurderoy, R. (2001). Age, Order of Entry, Strategic Orientation and Organizational Performance. *Journal of Business Venturing* 16: 471–494.
- Du Rietz, A. and Henrekson, M. (2000). Testing the Female Underperformance Hypothesis. *Small Business Economics* 14 (1): 1–10.
- Ensley, M., Pearson, A. and Amason, A. (2002). Understanding the Dynamics of New Venture Top Management Teams—Cohesion, Conflict and New Venture Performance. *Journal of Business Venturing* 17 (4): 365–386.

- Evans, D. (1987). The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries. *Journal of Industrial Economics* 35: 567–581.
- Evans, D. (1987). Tests of Alternative Theories of Firm Growth. *Journal of Political Economy* 95 (4): 657–674.
- Evans, D. and Leighton, L. (1999). Some Empirical Aspects of Entrepreneurship. *American Economic Review* 79: 519–535.
- Feindt, S., Jeffcoate, J. and Chappell, C. (2002). Identifying Success Factors for Rapid Growth in SME E-Commerce. *Small Business Economics* 19 (1): 51–62.
- Folster, S. (2000). Do Entrepreneurs Create Jobs? *Small Business Economics* 14 (2): 137–148.
- Harding, R., Cowling, M. and Turner, N. (2003). *The Missing Link: From Performance to Productivity*. The Work Foundation, London (<http://www.theworkfoundation.com/research/work>).
- Heshmati, A. (2000). On the Growth of Micro and Small Firms: Evidence from Sweden. *Small Business Economics* 17 (3): 213–228.
- Jovanovic, B. (1982). Selection and Evolution of Industry. *Econometrica* 50: 649–670.
- Lerner, M. and Haber, S. (2000). Performance Factors of Small Tourism Ventures: The Interface of Tourism, Entrepreneurship and the Environment. *Journal of Business Venturing* 16: 77–100.
- Lin, Z., Picot, G. and Compton, J. (2000). The Entry and Exit Dynamics of Self-Employment in Canada. *Small Business Economics* 15 (2): 105–125.
- McGee, J., Dowling, M. and Megginson, W. (1995). Cooperative Strategy and New Venture Performance: The Role of Business Strategy and Management. *Strategic Management Journal* 16 (7): 565–580.
- Meager, N., Bates, P. and Cowling, M. (2003). An Evaluation of Business Start-up Support for Young People. *National Institute Economic Review* 186 (4): 59–72.
- Meager, N. (1996). From Unemployment to Self-Employment: Labor Market Policies for Business Start-Up. In *International Handbook for Labor Market Policy Evaluation*, G. Schmid, J. O'Reilly and K. Schomann (Eds). Edward Elgar, Cheltenham.
- OECD (1996). *Job Creation and Loss: Analysis, Policy and Development*. OECD, Paris.
- Pennings, J., Lee, K. and van Witteloostuijn, A. (1998). Human Capital, Social Capital and Firm Dissolution. *Academy of Management Journal* 41 (4): 425–440.
- Phillips, B. and Kirchoff, B. (1989). Formation, Growth and Survival: Small Firm Dynamics in the U.S. Economy. *Small Business Economics* 1: 65–74.
- Pfeiffer, F. and Reize, F. (2000). Business Start-Ups by the Unemployed. *Labor Economics* 7 (5): 629–663.
- Prais, S. (1976). *The Evolution of Giant Firms in the U.K.* Cambridge University Press, Cambridge.
- Reid, G. (1995). Early Life Cycle Behavior of Micro Firms in Scotland. *Small Business Economics* 7: 89–95.
- Reid, G. and Smith, J. (2000). What Makes a New Business Start-Up Successful? *Small Business Economics* 14: 165–182.
- Robbins, D., Pantuosco, L., Parker, D. and Fuller, B. (2000). An Empirical Assessment of the Contribution of Small Business Employment to U.S. State Economic Performance. *Small Business Economics* 15 (4): 293–302.
- Robinson, K. and McGougall, P. (2001). Entry Barriers and New Venture Performance. *Strategic Management Journal* 22: 659–685.
- Robson, P. and Bennett, R. (2000). SME Growth: The Relationship with Business Advice and External Collaboration. *Small Business Economics* 15 (3): 193–208.
- Samuels, J. (1965). Size and the Growth of Firms. *Review of Economic Studies* 32: 105–112.
- Sapienza, H. and Grimm, C. (1997). Founder Characteristics, Start-Up Process, and Strategy/Structure Variables as Predictors of Shortline Railroad Performance. *Entrepreneurship, Theory & Practice* 22: 5–24.

- Schumpeter, J. (1942). *Capitalism, Socialism and Democracy*. Harper and Row, New York.
- Scott, M. and Bruce, R. (1987). Five Stages of Growth in Small Businesses. *Long Range Planning* 20: 45–52.
- Storey, D. (1994). *Understanding the Small Business Sector*. Routledge, London.
- Taylor, M. (1999). Survival of the Fittest? An Analysis of Self-employment in Britain. *Economic Journal* 109: 14–55.
- Van Gelderen, M., Frese, M. and Thurik, R. (2000). Strategies, Uncertainty and Performance of Small Business Start-Ups. *Small Business Economics* 15: 165–181.
- Van Praag, M. (1994). *Determinants of Self-Employment Duration: A Competing Risks Model*. Mimeo, Tinbergen Institute, University of Amsterdam, Amsterdam.
- Wesson, T. and de Figueiredo, J. (2001). The Importance of Focus to Market Entrants: A Study of Microbrewery Performance. *Journal of Business Venturing* 16 (4): 377–403.
- Westhead, P., Ucbasaran, D. and Wright, M. (2003). Differences between Private Firms Owned by Novice, Serial and Portfolio Entrepreneurs. *Regional Studies* 37 (2): 187–200.
- Westhead, P. and Cowling, M. (1995). Employment Change in Independent Owner-Managed High Technology Firms in Great Britain. *Small Business Economics* 7 (2): 111–140.
- Zahra, S. (1996). Technology Strategy and New Venture Performance: A Study of Corporate Sponsored and Independent Biotechnology Ventures. *Journal of Business Venturing* 11 (4): 289–321.

C. MIRJAM VAN PRAAG

*University of Amsterdam and Tinbergen Institute, Roetersstraat 11, 1018WB Amsterdam,  
The Netherlands*

*E-mail: c.m.vanpraag@uva.nl*

## 17. Venture Performance and Venture Inputs: The Role of Human and Financial Capital

### 1. INTRODUCTION

No matter how important it is to understand the issues and determinants of venture *creation*, the inevitable next question concerns venture *performance*, the topic of this chapter. Business development is a stressful and highly demanding type of activity which often ends in failure: 20% of new ventures do not survive their first year (Fritsch et al., 2004). As Audretsch and Keilbach (2003) indicated, the social costs of business failure are immense, whereas the social benefits of successful enterprising are enormous in terms of innovation, growth and competition. Hence, devising policy measures to lower entry barriers for potentially successful starters and to increase useful support can have a large impact on the economic development of countries. This is why many countries and international organizations like the European Commission, IMF and The World Bank put energy into designing and implementing such measures.

The effective design of policy measures to create successful entrepreneurship requires quantitative knowledge about the determinants of venture performance. This chapter describes the history as well as the state of the art of research into these determinants and reviews the available evidence. In so doing, the chapter focuses on two main determinants of venture performance: human capital and financial capital.

The motivation for this focus is fourfold. First, broad research, both historically and contemporary, shows that these two factors are the main individual drivers of venture performance (and also of venture development; cf. Le, 1999; Van Praag, 1999, 2003; and Chapter 15 of this volume). In general, they are more influential for performance than, for instance, ethnicity,

family background or the business strategy and organization of the small business starter. Second, the relationship between performance and some of the aforementioned potential determinants, such as an entrepreneur's business strategy, is, in turn, determined by the human and financial capital of the small business owner. Third, policy interventions relating to these two determinants are relatively straightforward since both human and financial capital result from individual decisions that are shaped by the political and business environments. Finally, the state of the art of the research into these two factors is most developed, though certainly not without critique.

Nevertheless, the focus on human and financial capital might limit the interest of the contribution of the chapter. To further position this chapter in the entrepreneurship research area, the perspective is economic rather than psychological, sociological or business organizational of character. I do discuss the (human and financial capital) investments that entrepreneurs make and the impact of these investments on the performance of their businesses. I do not discuss psychological traits that might affect business success, such as locus-of-control beliefs or risk taking propensity (cf. Parker, 2004, Chapter 3). Though interesting *per se*, the policy relevance of such determinants of entrepreneurial performance is limited as these characteristics of individuals are scarcely amenable to policy intervention. I neither discuss sociological issues such as the intergenerational transfer of business acumen nor the effect of parental background more generally, nor the effects of race, gender and/or discrimination on business performance. An overview of this literature can be found in Parker (2004, Chapters 3 and 4). Empirical regularities about business strategies of entrepreneurs that lead to business success are also lacking from this overview. That is a separate and rather case-study-based literature which goes beyond the scope of this chapter. The economic perspective that I employ is microeconomic, not macroeconomic. Although highly policy relevant, I do not discuss the effects on venture performance of economic circumstances such as the unemployment rate or the growth of the economy (see Parker, 2004, Section 3.3, for an overview). I do discuss the effects of an individual's choice to invest in human and (access to) financial capital on venture performance.

The remainder of this chapter is organized as follows. Section 2 discusses empirical evidence of the relationship between venture performance and human capital. In so doing, the focus is on the most prominent manifestation of human capital: education. The other important determinant of human capital, that is, labor experience, is outside the scope of this chapter. Research into the role of various sorts of labor experience, such as general labor experience, within-industry experience, management or entrepreneur experience has delivered scant but interesting results (Parker, 2004, Chapter 3, p. 72). Section 3 is devoted to the relationship between venture performance and financial capital (constraints). Both Sections 2 and 3 start with economic theory

about the relationship and then discuss the empirical evidence. Subsequently, both sections describe a recent study using an improved empirical strategy. They then discuss remaining shortcomings in the empirical literature. Section 4 concludes.

## 2. PERFORMANCE AND HUMAN CAPITAL<sup>1</sup>

### 2.1. Theory

Theories about the relationship between education and successful entrepreneurship have a long history in economics. The neoclassical economist Alfred Marshall was one of the first to express a view on this relationship in 1890 in his famous book *Principles of Economics*. His narrative theory poses that successful entrepreneurship requires, besides being a “natural leader of men” (Marshall, 1890[1930], p. 298), specialized and general abilities. Specialized abilities such as knowledge of the trade, power of forecasting, of seeing business opportunities and of undertaking risks is obtained through experience. General ability, on the other hand, such as “to be able to bear in mind many things at a time, to have everything ready when wanted, to act promptly, and show resource when anything goes wrong, to accommodate oneself quickly to changes. . .” (ibid, p. 206–207), depends on family background, innate ability, and . . . education.

The later neoclassical tradition, which has been based on formal modeling and empirical testing, has for a long time not paid much attention to the entrepreneur, though it has contributed importantly to the development of human capital theory. “Since the so-called human capital revolution of the 1960s and 1970s, a large body of both theoretical and empirical research has emerged, in which the decision to invest in human capital and the relations between education and earnings are explored” (Oosterbeek, 1992, p. 1).

Human capital theory in general indicates that previous knowledge plays a critical role in intellectual performance; it assists in the integration and accumulation of new knowledge as well as the adaptation to new situations (Weick, 1996). Knowledge may be defined as either tacit (“know-how”) or explicit (“know what”) (Davidsson and Honig, 2003). Individuals may increase their knowledge through formal education while informal education is gained through work or “life” experience. Indeed, the main factors affecting earnings are education and experience in the Mincerian specification (cf. Mincer, 1974). This section focuses on the effect of formal education on the entrepreneur’s performance.

Tests of human capital theory have mostly been performed on employees. Whether the productive effect of education is the same, larger, or smaller for entrepreneurs is an empirical matter. On the one hand, there is no reason

to believe *a priori* that the same qualitative relationship would not hold for the entrepreneurial sector of the labor market. As Davidsson and Honig (2003) asserted, making entrepreneurial decisions about complex problems utilizes an interaction of tacit and explicit knowledge. On the other hand, as asserted by Parker (2004, Chapter 1) there is reason to presume that the magnitude of the productive effect is smaller for entrepreneurs than for employees. Entrepreneurial success is likely to depend, more so than for employees, on numerous factors other than education, implying that the productive effect of education will be smaller for entrepreneurs. Formal education might even cause habits, attitudes or preferences that are counterproductive for entrepreneurial endeavor.

Human capital is not only acknowledged for its productive effect on labor supply, it also has value as a signal of productive ability in labor markets without complete information (Spence, 1973; Wolpin, 1977; Riley, 2002).<sup>2</sup> Like the empirical validity of human capital theory, the empirical validity of the screening value of education has also largely been assessed on the employees' subsample of the labor market. The question is whether the signaling effect, next to a productive effect of education, is as likely for entrepreneurs as for employees.

Many of the empirical tests devised to quantify the signaling effect of education for employees (Wolpin, 1977) assume that entrepreneurs, not having a prospective employer, can be treated as an unscreened control group. An empirical test in support of the screening hypothesis would therefore demonstrate that entrepreneurs have a smaller positive return to schooling than employees. I question the assumption that such a signal would be useless for entrepreneurs for two reasons. First, when acquiring education, the future entrepreneur might intend to work for an employer first. Second, there might be substantial screening from prospective capital suppliers, customers and other stakeholders. Education is then used as a signal. The returns to education (RTE) could thus be of similar levels for employees and entrepreneurs. A comparison of the returns to education, both as a result of increased productivity and as a result of its signaling value, is therefore an empirical matter.

## 2.2. Empirical Evidence

The relationship between schooling and entrepreneurship outcomes has been measured in various empirical studies. Van der Sluis, Van Praag and Vijverberg (2003) (VVV hereafter) provide an overview of such empirical studies into the impact of schooling on entrepreneurship selection and performance. They perform an extensive meta-analysis to assess whether there are any consistent findings from the vast empirical literature about the impact of education on entrepreneurship in industrialized countries.<sup>3</sup> To what extent are



performance and education related? And, is the beneficial effect of education on performance stronger or weaker for entrepreneurs than for employees?

Five important outcomes emerge from this meta-analysis. First, most researchers have focused on finding a quantitative measure of the effect of education levels rather than types. Education levels are measured either in terms of the number of years of formal education or in terms of the highest schooling level attained. Until recently, little effort has been put forth to differentiate the effect on entrepreneurship performance of various types of education. However, recently, various interesting turns have been taken by a handful of researchers in this area, that is, Lazaer (2004), Silva (2004) and Wagner (2002). They perform analyses to answer the question: "Do people require a one-dimensional specialist type of education to become successful as entrepreneurs? Or do they need the requirements of 'Jacks-of-all-Trades' for being successful as entrepreneurs?" If anything, the successful entrepreneur appears to be a Jack-of-all-Trades, rather than a specialist. Further questions, such as the value of vocational schooling relative to general education tracks and the value of specific entrepreneurship oriented courses have not yet been measured adequately (see Chapter 4 of this volume).

Second, the impact of an individual's schooling level on performance is significantly positive for 67% of the studies included in the meta-analysis. Performance has been measured as entrepreneurial income, profit, survival probabilities or, in terms of growth of for instance personnel, sales or profit. We conclude that entrepreneurship performance is significantly affected by schooling.

Third, the meta-analysis gives insight into the level of the RTE for entrepreneurs. This insight, though, can only be based on a small sub-sample of U.S. observations that uses similar measures for education and earnings. The return to a marginal year of schooling in terms of the income it generates is 6.1% on average.

Fourth, the meta-analysis allows a comparison of the rate of return to education for entrepreneurs to the returns for employees. This comparison is based on the results from 20 studies that compare the two groups of labor market participants using one dataset and thereby one set of definitions, time period, country and the like. From these studies the fourth result is obtained: the RTE are of similar levels for employees and entrepreneurs. More specifically, all studies pertaining to Europe indicate that the RTE are slightly lower for entrepreneurs than for employees. However, the opposite result is found for the studies that pertain to the United States.

The fifth conclusion from the meta-analysis is rather striking: most studies measuring the quantitative effect of years of education on performance have merely measured the (conditional) correlation between education and

performance by means of OLS rather than the causal effect, the estimate of interest.

There are at least two possible sources of inconsistency when OLS is used to estimate this relationship. First, the schooling decision is probably endogenous in a performance equation because individuals are likely to base their schooling investment decision, at least in part, on the expected payoffs to their investment. Second, there may be unobserved individual characteristics, such as ability and motivation, that affect both the schooling level attained and subsequent business performance. The omission of these unobserved characteristics from a performance equation would also serve to bias OLS estimates. Several methods to cope with these problems have been applied recently to estimate the RTE for employees. The general conclusion is that OLS estimates of the RTE for employees are biased downward (Ashenfelter et al., 1999).

The potential bias also raises suspicion about the comparisons of RTEs for entrepreneurs and employees. The neglect of unobserved influential characteristics and not dealing with the endogenous nature of the education decision can have a different impact on the estimate of the RTE for entrepreneurs and employees; cf. Griliches (1977). As a result, the conclusions from such comparisons should be re-evaluated.

The latter conclusion from the meta-analysis, that is, that the effect of education on entrepreneur performance has not yet been measured properly, puts all other conclusions in a different perspective. The remainder of this section is devoted to a short presentation of possible methods to obtain more consistent estimates and to a brief discussion of a recent application of such a method. This will be followed by a re-evaluation of the conclusions from the meta-analysis.

### *2.3. Measuring the Effect of Education on Outcomes*

There are basically four methods to account for the potential problems of endogeneity and/or unobserved heterogeneity when estimating the RTE. All four have been applied to the estimation of the RTE for employees (Ashenfelter et al., 1999).

The first strategy to cope with unobserved ability is trying to make the unobservable observable. Various proxies of intelligence and test scores have been added to equations from which estimates of RTE result. The effects so far of adding ability controls on the estimated returns to education are ambiguous (see Ashenfelter et al., 1999, Table 3).

The second strategy to identify causal effects involves setting up a randomized experiment. This approach has not yet been much applied in labor economics research (Leuven et al., 2003). The proper design of an experiment requires a random assignment of individuals into a treatment group

(participating) and a control group (not participating). In this manner, the choice to follow education is “forced.” The problem is that setting up an experiment where some people do not get (higher) education but others do is not ethically feasible.

The third strategy uses the variation in schooling, and income between monozygotic twins to estimate RTE. This approach has been used to identify employees’ RTE (e.g., Ashenfelter and Krueger, 1994, and Bonjour et al., 2003). The basic idea is that monozygotic twins share the same genetic endowment and usually experience even more similar environments than non-twin siblings. Thus, comparing monozygotic twins should control thoroughly for otherwise unobserved heterogeneity. In general, these studies render a higher estimate of the RTE of employees than OLS (Ashenfelter et al., 1999).

The fourth strategy identifies causal effects using an instrumental variable (IV) approach. The idea is to imitate a field experiment where economic characteristics are randomly allocated among individuals to estimate their effects. This strategy therefore enables the measurement of the effect of schooling, assuming a random allocation of schooling levels among individuals, independent of their expected pay-offs or relevant unobserved background variables. In general, IV estimates of the returns to an employee’s education are higher than OLS estimates (Ashenfelter et al., 1999). In the following, I shall discuss an application of the IV identification strategy to estimate the returns to education for entrepreneurs as well as employees.

#### *2.4. Application: Returns to Education for Entrepreneurs and Employees in the U.S.*

Van der Sluis, Van Praag and Van Witteloostuijn (2004) (VW hereafter) compare the magnitude of the returns to education for entrepreneurs and employees. Using the same methodology for both samples, drawn from the National Longitudinal Survey of Youth (NLSY), they estimate a random effects model using IV while including a set of detailed ability proxies. The sample VW use from the NLSY is a rich panel consisting of more than 6000 individuals and 19 annual waves.

An important feature of the NLSY is the presence of detailed family background variables. VW use some of these variables as identifying instruments for the respondent’s education. These family background characteristics are possibly good predictors of the educational level of the respondent while otherwise independent of their future wage. Blackburn and Neumark (1993) apply the same methodology on the same sample to estimate the returns to education for employees.

A second relevant feature of the sample is that it includes both entrepreneurs and employees and records individuals’ switches between these states

over time. All entrepreneurship spells, also short ones, are recorded. Therefore, the subsample of entrepreneurs does not suffer from survival bias, that is, the RTE will not pertain to surviving entrepreneurs only. Moreover, the incomes and all other relevant variables are measured in a comparable way for both groups such that the RTE for employees and entrepreneurs can be estimated in a comparable fashion.

The results reveal that the OLS-estimate of the RTE is around 7% for both entrepreneurs and employees. In accordance with previous studies using U.S. data the returns are slightly higher for entrepreneurs than for employees, though of the same order of magnitude (Fredland and Little, 1981; Tucker, 1985, 1987; Evans and Leighton, 1990; Robinson and Sexton, 1994).

The estimation results when using an IV-approach are different and novel, giving significantly higher estimates of the RTE. The increase from 6.7% to 10.7% for employees is comparable to increases resulting from applying IV instead of OLS in previous applications, such as Blackburn and Neumark (1993). A novel observation is the even greater jump in the estimates pertaining to entrepreneurs: the IV estimate is twice as high as the OLS estimate of 7.1% and amounts to 14.2%. This leads to a remarkable result: the RTE for entrepreneurs are estimated to be significantly higher than for employees in the US. Previous research based on OLS estimates resulted in much smaller and insignificant differences. VVW performed various checks to assess the credibility and robustness of the result. They found that the result is due neither to selectivity (i.e., caused by different people choosing for entrepreneurship instead of wage employment) nor to a higher required *risk premium* obtained by higher educated entrepreneurs. So why is education more valuable for entrepreneurs?

VVW proposed a simple explanation: Entrepreneurs have more freedom to optimize their use of education. Entrepreneurs are not constrained by rules from superiors and can decide on how to put their education to its most productive use. This difference in ability to optimize the productivity of education for entrepreneurs and employees might therefore be an explanation for the higher returns to education for entrepreneurs.

## 2.5. Discussion

A meta-analysis of the empirical evidence on the relationship between education and entrepreneurship performance showed the following empirical regularities concerning entrepreneurship success and education *level*: (i) RTE is positive for entrepreneurs, (ii) RTE is 6%, (iii) the RTE is of similar size for entrepreneurs and for employees, and (iv) the empirical evidence may have produced (probably downward) biased results due to using estimation methods that fail to identify causal effects. The conclusions from the meta-

analysis should therefore be re-evaluated based on novel findings from a recent application of the IV methodology.

VVW have performed this study and find that (U.S.) entrepreneurs benefit more from an additional year of education than their employed counterpart: The RTE are shown to be higher for entrepreneurs (14% and 10%, respectively). As the difference in estimated returns is much smaller and insignificant when using OLS, I conclude that (but do not yet understand why) the bias resulting from neglecting endogeneity is larger for entrepreneurs than for employees.

Based on these results, I re-evaluate the conclusions from the meta-analysis and conclude as follows: (i) the RTE for entrepreneurs are indeed positive and (ii) much higher than previous measures indicated.: they seem to amount to 14% in the U.S., and (iii) are thereby significantly higher than the RTE for employees in the same country; (iv) previous (OLS) estimates were indeed biased.

However, we cannot draw strong conclusions yet. Replicating the VVW study with different data, also for various countries and possibly with different sets of instruments, would be useful to confirm the findings that are now based on one study only. Parker and Van Praag (2004) have recently made a step forward in this respect: They estimate the returns to education for Dutch entrepreneurs using a similar approach as VVW. They also find that the returns to education are higher for entrepreneurs than the percentage gains from education that are usually found for Dutch employees. Alternatively, these findings could be validated (or not) using different identification strategies, like twins research. In any case, more insight into the effect of education on the performance of entrepreneurs is required. Moreover, a deeper understanding of the types of education that provide the best basis for successful entrepreneurial endeavor should result from continuing on the interesting research routes set out by Lazear (2004) and others.

### 3. PERFORMANCE AND FINANCIAL CAPITAL

Government spending to increase the numbers of higher qualified entrepreneurs is not only explained by the social benefit pertaining to entrepreneurial endeavor, but also by the perceived existence of undesirable impediments to the supply of entrepreneurs. A lack of capital is one of these factors and is the focus of this section.

The objective of this section is to answer the question: To what extent is the performance of a small business founder's entrepreneurial venture, once started, affected by capital constraints at the time of inception? What happens to performance when an entrepreneur has insufficient capital to reach the optimal investment level or timing? Financial capital constraints might

prevent entrepreneurs from creating buffers against random shocks, thereby affecting the timing of investments negatively. Moreover, capital constraints might debar entrepreneurs from the pursuit of more capital-intensive strategies. Thus, what we are aiming at is measuring the effect of initial capital constraints on venture performance. Merely measuring the correlation between capital constraints and performance would not be sufficient, since it would (wrongly) include *spurious* factors that affect access to capital as well as performance directly such as ability and motivation. The distinction between causal and spurious factors is crucial since policy implications diverge. In the first case, supplying more capital to constrained entrepreneurs with suboptimal capital stocks would improve performance. In the second case, it will not because the capital constraint itself is not the binding restriction, but the factors underlying it are.

Much (empirical) research effort has been put into measuring the effect of capital constraints on the *selection* of individuals into entrepreneurial positions.<sup>4</sup> The general conclusion is that capital constraints bind. A significant proportion of individuals willing to enter entrepreneurship are hampered by a lack of sufficient capital. Blanchflower and Oswald (1998) presented recent survey evidence that many individuals who are currently employees would prefer to be self-employed, thereby indicating the existence of impediments at large. The International Social Survey Programme of 1989 asked random sets of individuals from 11 industrialized countries the question: "Suppose you were working and could choose between different kinds of jobs. Which of the following would you choose? I would choose (a) being an employee; (b) being self-employed; (c) can't choose." It turned out that an astonishing 63% of Americans, 48% of British and 49% of Germans opted for answer (b), whereas the actual proportion of self-employed in these countries is at most 15%. Indeed, as arises from the same article, self-employed people appear to be much more (work-) satisfied than employees, despite the latter earning significantly higher and more secured incomes (Hamilton, 2000). Blanchflower and Oswald provided evidence that the significant impediments indicated by these results are, in 50% of cases, due to lack of capital. Capital markets are neither efficient nor market clearing for the segment of new firms (Fazzari, 1988). Personal savings and loans from friends and relatives is by far the largest source of capital in newly started firms (e.g., Parker, 2004, p. 137). Bank loans, often highly collateralized, are the second most common source of capital for new firms (e.g., Astebro and Bernhardt, 2003; Parker, 2004, p. 137; and see Chapter 8 of this volume). This order of preference, that is, internal funds, external loans and then external asset-based finance, is in accordance with the "pecking order theory" (Myers and Majluf, 1984; Cosh et al., 2005).

The common theoretical explanation for credit rationing *vis-à-vis* newly founded firms is a severe lack of observable and verifiable information about

the entrepreneur's type, their plans and the risk associated with these plans. Moreover, the entrepreneur, financing a venture by means of borrowed capital, might have intentions that conflict with these of the supplier of the loan. The asymmetry of information about the entrepreneur's type and behavior will potentially lead to agency problems: adverse selection and moral hazard (LeRoy and Singell, 1987; Boadway et al., 1998; De Meza and Webb, 2000). The foresight of these problems prevents the start of a significant proportion of ventures. A negative correlation results between access to capital and entrepreneurship entry.

Research effort has also been devoted, though to a lesser extent, to measuring the correlation between access to capital and entrepreneurship performance once the stage of startup has been successfully completed.<sup>5</sup> This section aims to contribute to this category of research.

### 3.1. Theory

A lively theoretical debate has existed about the relationship between access to capital and investment decisions of entrepreneurs ever since entrepreneurship has become a topic of study. The first stream of thought assumes capital markets to be perfect. External funds provide a perfect substitute for internal capital in this full information case. An entrepreneur's financial conditions are irrelevant to investment. Investment decisions are independent of whether one needs to "pay" the opportunity cost of capital ownership or the interest rate of borrowing money. Proponents of this view can be traced back to Richard Cantillon (1979; original edition 1755). Cantillon implicitly assumed perfect(ly accessible) capital markets. Later influential economists agreeing with this view of the independence of capital from investment decisions and entrepreneurial performance included Schumpeter (1934; original edition 1911) and Kirzner (1973).

The second stream of research in entrepreneurship assumes imperfect capital markets due to the existence of imperfect and asymmetric information. The latter makes it very costly and sometimes even impossible for providers of external finance to evaluate the quality of an entrepreneur's investment opportunities. This might debar (some) entrepreneurs from sufficient access to external capital. As a consequence, internal and external capital sources are not perfectly substitutable. This view has a history in economic thought of entrepreneurship, too. The performance of the entrepreneur in the classical and neoclassical theories of Say (1971; original edition 1803) and Marshall (1930; original edition 1890), respectively, is hindered by a lack of own capital since borrowed capital requires a reputation (Say) or a risk premium (Marshall). Knight (1971; original edition 1921) held the same view. According to these

views, investment decisions may depend on capital ownership and capital constraints might bind.

Interestingly, before this debate regained interest among entrepreneurship researchers in the 1980s, a similar debate took place among researchers in the area of corporate finance and investment (all company sizes included) in the late 1950s and early 1960s; see, for example, Meyer and Kuh (1957). However, since the middle 1960s, most research in this area has isolated real firm (investment) decisions from purely financing factors. Modigliani and Miller (1958) provided the theoretical basis for this second school of thought by demonstrating the irrelevance of financial structure and policy for real investments (under certain conditions). Their key insight was that a firm's financial structure (read: entrepreneur's own wealth) will not affect its market value (read: the entrepreneur's firm performance) under certain conditions. Applied to a firm's investment decisions, this finding by Modigliani and Miller provided a foundation for the neoclassical theory of investment in which firms are assumed to face costs of capital that do not depend on the firm's particular financial structure. However, Fazzari (1988) established convincing empirical evidence for the existence of a capital constraint, especially in the small firm case. Access to external (borrowed) capital is difficult or at least more costly than is investment by means of internal capital.

### *3.2. Empirical Evidence: Capital Constraints and Performance*

The continuation of the debate in the entrepreneurship research area, starting in the late 1980s, had a large empirical component. Most empirical research is based on the view, as expressed by De Meza and Webb (2000), that asymmetric information applies at least to the type of entrepreneur (hidden type problem potentially leading to adverse selection) but may extend to the behavior of the entrepreneur (hidden action problem potentially leading to moral hazard). Furthermore, to prevent adverse selection in the credit market, the point of departure is not credit rationing in response to the hidden type problem but credit scoring instead. Credit scoring, as applied in almost all real-life cases (De Meza and Webb, 2000) involves suppliers of money using some selection procedures based on a set of indicator variables for the expected performance and risk of entrepreneurs and their projects. Those failing to score sufficiently high on the criteria used are denied credit for whatever interest rate they might be willing to pay. As a consequence, several indicators of entrepreneurship performance such as education and experience might turn out to be indicators of access to capital (Bates, 1990; Scherr et al., 1993). This clarifies part of the discussion below about the inclusion of human (sometimes also social) capital variables in empirical models. To discriminate between the full information and asymmetric information case, several categories of empirical research have been



TABLE 17-1 Evidence about the effects of financial capital on performance

| Lack of access to capital measure/<br>performance measure | Assets                    | Inheritance    | Windfall<br>gains |
|---|---------------------------|----------------|-------------------|
| Earnings  | EJ: +                     | HJR: +         |                   |
| Survival  | CGW: +; T: +; vP: 0; C: 0 | HJR: +; BFN: + | LO: +             |
| Growth  | CGW: +; CTM: 0            | BFN: +; CTM: + | CTM: 0            |

BFN: Burke, FitzRoy and Nolan (2002); C: Cressy (1996); CGW: Cooper, Gimeno-Gascon and Woo (1994); CTM: Cowling, Taylor and Mitchell (2004); EJ: Evans and Jovanovic (1989); HJR: Holtz-Eakin, Joulfaian and Rosen (1994); LO: Lindh and Ohlsson (1996); T: Taylor (1999); vP: Van Praag (2003).

performed.<sup>6</sup> An overview is given in Table 17-1. The entries in the table show which studies have used a particular measure of capital constraints (columns) in combination with a particular performance measure (rows). The following subsections discuss each column of the table in turn.

*3.2.1. Relationship Between Assets and Performance* Many researchers have related the size of family assets to earnings from (or job creation, growth or survival of) entrepreneurial ventures. Both Evans and Jovanovic (1989) (henceforth EJ) and Cooper et al. (1994) (henceforth CGW) found a positive association between assets and performance for U.S. entrepreneurs. Taylor’s (1999) result pertaining to the U.K. is supportive of EJ and CGW. The effect of a dummy indicating whether the respondent had received interest or dividend payments exceeding £100 is negative on the hazard and thus has a positive effect on survival. Van Praag (2003) also related financial variables, that is, assets and a dummy for home ownership (frequently used as collateral), to survival of young entrepreneurs in the U.S. The effect of these variables on the hazard out of entrepreneurship is insignificant. Cressy’s (1996) insignificant result on survival for the U.K. supports Van Praag’s finding. Furthermore, Cowling et al. (2004) estimated that assets do not increase job creation by British entrepreneurs while controlling for inheritance receipts and windfall gains.

Several drawbacks are attached to the studies in this category. First, the possibility of obtaining external finance remains unconsidered. It is assumed that the “external route to obtain finance” is totally inaccessible. Second, a monotone relationship is assumed between assets and performance, while in reality it might well be the case that up to a certain point more access to capital might help in enhancing performance, but eventually “enough is enough.” This possible discontinuity in the relationship is not taken into account in this approach. A third drawback of the method in general is that “family assets” is not an exogenous variable. Without binding capital constraints, a correlation could still exist between assets and performance because of the entrepreneur’s ability (“earning power”) affecting both quantities. A fourth drawback, finally,

is that assets in general are badly reported in individual survey research and therefore unreliable figures plagued with measurement error.<sup>7</sup>

*3.2.2. Relationship Between Inheritance Receipt and Performance* One of the major drawbacks of the approach of relating assets, as a measure of access to capital, to new venture performance is the possible endogenous character of assets. An interesting alternative indicator might be the receipt of an inheritance: “The receipt of an inheritance is about as close to a ‘natural experiment’ as one is likely to get in this area, thereby reducing potential endogeneity problems” (Blanchflower and Oswald, 1998). Holtz-Eakin et al. (1994a) (HJR henceforth) were the first to estimate the relationship of this inheritance variable with firm *performance* rather than entry. They found a positive effect from receiving an inheritance on firm survival and earnings in the U.S. Burke et al. (2002) estimated the effect of inheritances on both entry and performance where the latter is measured as survival and employment growth. They found all these relationships to be significantly positive. Cowling et al. (2004) found a positive effect of inheritances on job creation by entrepreneurs.

This innovative approach however only solves the third of the four drawbacks attached to the first approach. An additional disadvantage is evoked by the inheritance approach: “We find that young men’s own financial assets exert a statistically significant but quantitatively modest effect on the transition to self-employment. In contrast, the capital of parents exerts a large influence. Parents’ strongest effect runs not through financial means, but rather through human capital, i.e. the intergenerational correlation in self-employment” (Dunn and Holtz-Eakin, 2000).<sup>8</sup>

*3.2.3. Relationship Between Windfall Gains and Performance* Lindh and Ohlsson (1996) estimated the effect of windfall gains on the probability of *being* self-employed on a sample in Sweden. They consider windfall gains as a dummy variable indicating whether people have ever won in lotteries or have ever obtained personal or spousal inheritances. They find significant effects on self-employment of both inheritances and lottery prizes. However, upon inclusion of additional control variables (human capital) the significant effect of inheritance receipts vanishes whereas the effect of lottery prizes remains significant. This supports the finding by Dunn and Holtz-Eakin (2000) about the intergenerational correlation of entrepreneurship. Cowlings et al. (2004), though, found a positive effect of inheritance receipts on job creation, but they did not find such an effect of alternative indicators of windfall gains. The windfall gains approach, ingenious though it is, does not solve the majority of the drawbacks associated with the first approach, though it solves the problem of endogeneity.<sup>9</sup>

The following model set-up clarifies the first two drawbacks of the existing estimation methods: (i) the possibility of obtaining external finance

remains unconsidered and (ii) a monotone relationship is assumed between assets and performance.<sup>10</sup>

### 3.3. Model Set-up

Measure entrepreneurial performance  $P_i$  in terms of gross receipts as in Holtz-Eakin et al. (1994a) and consistent with Evans and Jovanovic (1989):

$$P_i = \theta_i f(k_i)\varepsilon, \tag{1}$$

where  $\theta_i$  is individual  $i$ 's entrepreneurial ability or business acumen,  $f(\cdot)$  is a production function with one input, capital ( $k_i$ ), and  $\varepsilon$  is a random factor to the production process. Individuals know their ability, unlike the analyst or banker who observes an indicator function of ability,  $\tilde{\theta}_i$  only. Ability varies across individuals. It is assumed that  $\varepsilon$  has mean 1 and finite variance and that  $f(0) > 0$ . The firm can produce output even in the absence of any inputs, other than the entrepreneur's ability, as for example in the professional services industry.

$A_i$  is defined as the value of the individual's personal assets, hence  $A_i - k_i$  generates capital income at rate  $r$ . The (risk neutral) entrepreneur maximizes total income:

$$y_i = \theta_i f(k_i)\varepsilon + r(A_i - k_i). \tag{2}$$

The optimal investment level of capital into the venture is therefore defined by:

$$\theta_i f'(k_i^*) = r. \tag{3}$$

I assume that  $A_i$  is a nondecreasing function of  $\theta_i$ : Entrepreneurial ability is an indicator for general "earning power" from which assets might have resulted. The relationship between entrepreneurial ability and the amount of external capital required,  $k_i^* - A_i$ , is therefore ambiguous.

Access by individual entrepreneurs to the most desirable amount of external capital,  $l_i^* = k_i^* - A_i \geq 0$  at price  $r$  is constrained by the factor  $\beta_i$ , where  $0 \leq \beta_i \leq 1$ .  $\beta_i = 1$  represents the fully constrained entrepreneur;  $\beta_i = 0$  the unconstrained. The amount of external capital obtained is  $l_i = k_i - A_i = (1 - \beta_i)l_i^* = (1 - \beta_i)(k_i^* - A_i)$  for all entrepreneurs. The value of  $\beta_i$  depends on "borrowing power" which is dependent in turn on collateral and  $\tilde{\theta}_i$ .

The central question is to what extent  $\beta_i$  creates performance losses, that is, the effect of  $\beta_i$  on the expected (constrained) performance:

$$P_i = \theta_i f(k_i^* - \beta_i(k_i^* - A_i)) = \theta_i f(A_i + (1 - \beta_i)(k_i^* - A_i)). \tag{4}$$

In order to get rid of the intruding effect of ability on the relationship between absolute performance and capital constraints, relative performance is considered:

$$\log P_i = \log \theta_i + \log f(k_i^* - \beta_i(k_i^* - A_i)). \quad (5)$$

Equations (4) and (5) immediately show a drawback of the approaches discussed in the previous section. Simply looking at how a change in  $A_i$  affects performance does not measure the effect of capital constraints on performance.

In the following, I discuss a recent study by Bosma, Van Praag, and De Wit (2003) that attempts to measure the effect of capital constraints on performance while limiting as much as possible the biases resulting from the drawbacks pertaining to previous measurements.<sup>11</sup>

#### 3.4. Application: Capital Constraints and Performance

Bosma, Van Praag, and De Wit (2003) (BVD henceforth) evaluated the effect of capital constraints on entrepreneurial performance based on a panel of 1000 Dutch entrepreneurs. They found that initial capital constraints harm entrepreneurs' performance. They use a direct individual indicator variable for initial capital constraints so that policy implications will become more evident.

The empirical proxy for initial capital constraints,  $\beta_i$ , is a dummy variable formed by the answer to the question: "Did you experience problems in obtaining sufficient (external) capital at the start of your venture?" The distribution of answers in the sample is as follows:

|                                     |     |
|-------------------------------------|-----|
| Yes, and I didn't solve the problem | 7%  |
| Yes, but I solved the problem       | 17% |
| No                                  | 76% |

BVD considered the 7% of entrepreneurs who experienced these problems but did not solve them as being capital constrained ( $\beta_i = 1$ ).<sup>12</sup> The other 93% is characterized by  $\beta_i = 0$ . These entrepreneurs operate their businesses at the optimal level,  $k_i^*$ .

In this manner, BVD coped with the first two drawbacks attached to previously applied approaches. First, their estimate of  $\beta$ 's coefficient showed the effect on performance of being capital constrained for the group of entrepreneurs who are capital constrained. In contrast, "conventional" approaches merely generate an estimate of the effect of an increase in assets on performance. Second, their estimate of  $\beta$ 's coefficient embodies the effect of capital constraints that remain after the possibility of obtaining external finance has been explored. Other approaches assume that external finance is totally

inaccessible. Moreover, the fourth drawback, the issue that empirical measures of assets are plagued with measurement error, is also circumvented by not using such a measure. However, circumventing this measurement problem comes at a cost. BVD relied on self-reported subjective answers about capital constraints. Over- or under-reporting of this variable would lead to biased results.

However, a severe limitation of BVD's approach is that they did not solve the endogeneity issue, that is, the third drawback, although they try minimizing the bias in their estimates of  $\partial P_i / \partial \beta_i$  by controlling as much as possible for ability and motivation. They included in their regression analyses: (i) Various indicators of human and social capital that might affect lenders' decisions (as well as performance), thereby leading to a positive bias of the coefficient of interest. (ii) A vector of signals of entrepreneurial ability,  $\theta_i$ , based on the known result of credit scoring by external capital suppliers. In particular they considered the assignment of a loan by family/friends, banks, and by business partners as informative about unobserved heterogeneity. (iii) An indicator of time spent on other paid activities. Entrepreneurs spending more time on other paid activities will probably show weaker venture performance and simultaneously face lower capital constraints. Without any additional corrective measures, this spurious effect would be included in an estimate of the coefficient for  $\beta_i$  leading to a downward bias. (iv) Indicators for the amount of income generated from other paid activities and the (subjectively assessed) extent to which the entrepreneur is financially dependent on the venture's income. Financial independence from the venture might be a cause for lower capital constraints and might simultaneously result in a weaker motivation. Without correction, this spurious effect would again lead to a downward bias.<sup>13</sup>

The data they used result from questionnaires conducted on a representative sample of 1300 firm founders taken from all newly registered firms in the first quarter of 1994 with the Dutch Chamber of Commerce. The information from the 1994 questionnaire was used for the construction of potential determinants of performance. Entrepreneurial performance itself, measured by (the logarithm of) profits and survival duration, is exclusively measured by means of variables constructed from subsequent questionnaires in the years 1995–1997.<sup>14</sup> In this manner, problems of serially reversed causality are prevented.

The first half of Table 17-2 shows the result from the Tobit estimation with (log) profit as the dependent variable and the capital constraint as the independent variable. Each specification, going from left to right, adds a category of control variables. The estimation results are consistent with binding capital constraints. Entrepreneurs who suffer from a lack of capital for their initial business investments have 63% lower profits. As was expected, column II in Table 17-2 shows that the effect of capital constraints on profit diminishes

TABLE 17-2 *Estimation results: the effects of capital constraints on performance*

| Profit                                    | I        | II      | III     | IV      | V       | VI      |
|---|----------|---------|---------|---------|---------|---------|
| <i>Controls included</i>                  |          |         |         |         |         |         |
| Basic <sup>a</sup>                        | -0.63*** |         |         |         |         |         |
| Human capital controls <sup>b</sup>       |          | -0.59** |         |         |         |         |
| Social capital controls <sup>c</sup>      |          |         | -0.52** |         |         |         |
| Financial screening controls <sup>d</sup> |          |         |         | -0.51** |         |         |
| Time constraint control <sup>e</sup>      |          |         |         |         | -0.49** |         |
| Motivation controls <sup>f</sup>          |          |         |         |         |         | -0.51** |
| # obs.                                    | 1168     | 1168    | 1168    | 1168    | 1168    | 1168    |
| Log likelihood                            | -1643    | -1611   | -1599   | -1598   | -1595   | -1593   |
| <hr/>                                     |          |         |         |         |         |         |
| Duration                                  | I        | II      | III     | IV      | V       | VI      |
| <i>Controls included</i>                  |          |         |         |         |         |         |
| Basic <sup>a</sup>                        | -0.63*** |         |         |         |         |         |
| Human capital controls <sup>b</sup>       |          | -0.53** |         |         |         |         |
| Social capital controls <sup>c</sup>      |          |         | -0.47*  |         |         |         |
| Financial screening controls <sup>d</sup> |          |         |         | -0.47*  |         |         |
| Time constraint control <sup>e</sup>      |          |         |         |         | -0.47*  |         |
| Motivation controls <sup>f</sup>          |          |         |         |         |         | -0.48*  |
| Log likelihood                            |          | 1073    | 1073    | 1073    | 1073    | 1073    |
| # obs.                                    | 1168     | 1168    | 1168    | 1168    | 1168    | 1168    |
| Log likelihood                            | -1303    | -1285   | -1275   | -1275   | -1275   | -1275   |

Notes: <sup>a</sup>Basic controls include gender, hours worked at the start and a constant. <sup>b</sup>Human capital controls include age, age squared, employee experience, within-industry experience, experience in business ownership (dummy), experience relevant to business ownership (dummy) and education. <sup>c</sup>Social capital controls include the following dummies: (i) contact with entrepreneurs in a network; (ii) the utilization of various modes of information gathering; (iii) emotional support by a spouse; (iv) presence of a spouse. <sup>d</sup>Financial screening controls include the share of own capital in the business and the sources from which (co-)finance has been obtained. <sup>e</sup>The time constraint control is a dummy variable indicating whether one spends more than 20 hours per week on other paid activities. <sup>f</sup>The motivation controls include a dummy variable indicating whether other income is available and one indicating whether the entrepreneur is financially dependent on the business. \*Significance level  $p < 0.10$ ; \*\*Significance level  $p < 0.05$ .

(to 59%) when controlling for human capital effects, the capital constraint still being significant. Human capital, as was assumed, appears to simultaneously affect performance positively and the capital constraint negatively. The main factors of influence are various sorts of experience and education. Controlling for social capital factors (column III) also has a diminishing effect on the capital constraint. The coefficient decreases further from 59 to 52% and remains significant. The most important “social capital” factor is a spouse’s emotional support. Other social capital factors of influence are the exploitation of commercial contacts and contacts with fellow entrepreneurs. The correction for financial screening factors does not lead to a significant decrease of the

capital constraint (column IV). The capital constraint decreases from 52 to 51% only and remains significant. Moreover, financial screening factors have no additional significant effect on profits, suggesting that these factors do not reveal any heterogeneity in addition to human and social capital. The addition of the next two blocks of variables (columns V and VI in Table 17-2) aims at correcting the potential downward bias in the estimate for the capital constraint due to time and motivational constraints. It appears that the inclusion of such indicators does not increase the absolute value of the coefficient pertaining to the capital constraint. The remaining as “unbiased” as possible effect of the capital constraint on profit is a disadvantageous 51%.

The second half of Table 17-2 shows the effect of capital constraints on duration in the same manner. The effect of the capital constraint is of the same order of magnitude as in the profit equation, ranging from 63% without corrections to 48% percent with them. Column II shows that the inclusion of human capital factors diminishes the effect by 10 percentage points, whereas column III shows that social capital factors account for a decrease of another six percentage points. The other corrections have no significant effect. The remaining as “unbiased” as possible effect of the capital constraint on duration is a disadvantageous 48%.

The conclusion from Table 17-2 is that entrepreneurs who acknowledge unsolvable initial capital constraints experience lower profits, conditional upon survival, whereas their survival rate compares unfavorably to those who are not capital constrained. The size of the effect of capital constraints decreases when correcting for human and social capital factors, but it remains significant and relatively large. Financial screening, time and motivational constraints do not consistently show the expected effects, neither directly on performance, nor indirectly by changing the coefficient of the capital constraint. However, the direction of both the indirect and direct effects is as expected in all cases. I conclude that capital constraints apparently diminish investment opportunities in terms of size and/or timing.

### 3.5. Discussion

The theoretical debate about the relationship between financial capital constraints and entrepreneur performance has put forth two opposing views: (i) capital markets are perfect and therefore do not hinder entrepreneurs in their required investments with regards to the levels and timeliness; and (ii) capital markets do not supply the right amounts of capital to entrepreneurs due to asymmetric information. Empirical evidence has largely supported the second view. Capital constraints do appear to hinder entrepreneurial performance (see Table 17-1).

I have discussed the empirical strategies that have produced this evidence and their strengths and weaknesses. I then discussed a recent application of a different method to evaluate the effect of experiencing capital constraints on entrepreneurial performance. The study confirms that initial capital constraints and the implied suboptimal investment possibilities significantly hinder entrepreneurial performance.

However, the most important drawback of the discussed study is that it does not account for the endogenous character of the extent of capital constraints experienced by entrepreneurs in the entrepreneurial performance equation. This remains true no matter how many control variables are entered into the performance equation. None of the studies discussed in this section has yet accounted for this by means of instrumental variables or any of the other suitable approaches (discussed in Section 2.3). A recent study by Parker and Van Praag (2004) seems to be the exception. They estimated the effect of capital constraints on the incomes of Dutch entrepreneurs by using an instrumental variables approach while they accounted also for the endogenous nature of schooling. Their findings support those of BVD: capital constraints bind.

#### 4. CONCLUSION

In this chapter, I have discussed research into the effect of human and financial capital as venture inputs on venture performance. The results from such research might bear policy implications. Before focusing on policy, I will discuss the conclusions from the previous sections together with possible future avenues of research.

Based on a review of the literature, I argued in Section 2 that the returns to education for entrepreneurs need to be measured with the same methodological rigor as the studies on employees. The neglect of the endogenous nature of schooling is especially a problem and would possibly lead to underestimating the effect of education on performance. I discussed a recent study that has dealt with this problem. The result implies that previous estimates have indeed been biased downward, but does it shed new light on the conclusions from the literature with respect to the effect of education on venture performance (see Section 2.2)?

The first conclusion, that education has a significantly positive impact on entrepreneurs' performance, is supported. The second conclusion, that the average estimated rate of return to education for entrepreneurs of 6.1%, is not supported. The return to education for entrepreneurs turns out to be much higher, around 14%.

The third conclusion of the previous literature base, that the returns to education are slightly higher for entrepreneurs than for employees in the



United States (whereas the opposite applies for Europe), is not supported: The returns for entrepreneurs in the U.S. are shown to be much higher than the returns for employees (14% and 11%, respectively). This turnaround finding must be somewhat puzzling in light of the traditional studies that test screening hypotheses. Apparently entrepreneurs cannot be assumed to be an unscreened control group.

As for the effect of financial capital constraints on venture performance, the conclusion is that capital constraints lead to a suboptimal choice of investment opportunities and thereby to weaker venture performance. I argued in Section 3 that most previous studies have not actually measured the effect of capital constraints but rather the effect of assets or of an (random) increase in assets. The application discussed in Section 3 has been the first that measures the effect of capital constraints. They turn out to be binding.

Of course, the methods used are not without critique. We therefore argue that, in order to validate the results discussed, more analyses in this spirit should be performed. One other, but related, an issue of concern is the low “explanatory power” of equations explaining entrepreneurs’ incomes. VVW can explain only 28% of the variance in entrepreneurial income by the observed factors, compared to almost 50% in the employee part of the study. It is therefore possible that we are missing some important determinants of entrepreneurial performance. The full exploitation of human, social and financial capital as determinants of entrepreneur performance should therefore be considered. Education, for instance, has almost exclusively been defined in terms of the level of education. Future research should also focus on the specific direction and compilation of the education followed (vocational studies, technical studies, subjects studied or specific entrepreneurship orientated courses, etc.) following the route taken by Lazear (2004) and others. Moreover, the relative effect of the other determinant of human capital, labor experience and, in particular, of various sorts of labor market experience on business performance is an interesting area of research, too.

Concluding this discussion, the steps taken, especially by VVW, are a first contribution to the measurement of the causal and distinct effects of education and capital constraints on the performance of entrepreneurs. By using more recently developed estimation strategies, they found results that lead to different conclusions than previous research. However, these conclusions are tentative and more research in the same fashion is needed to validate the results on which they are based.

#### *4.1. Policy Implications*

Before I discuss the policy implications I elaborate on the assumptions required to be able to have those implications follow from the results. First,

I assume that the social return of entrepreneurial activity is larger than the private return that accrues to the entrepreneur herself. Second, I assume that the difference between the social and private benefits of entrepreneurial activity is larger than this difference is for employees. A successful entrepreneur is, for example, more likely to influence competition in a market than an employee. Moreover, entrepreneurs can usually bring new and innovative ideas more easily to the market than employees. Third, I assume that individuals invest in schooling at a stage in their lives at which they do not yet know, in general, whether they will become entrepreneurs or employees, or a (sequential) combination of both. As a consequence, investment in schooling is not motivated by the specific expected return derived from entrepreneurship but by some (weighted) average return of both employment modes. The fourth assumption is that (future) capital constraints do not affect the number of years of education pursued. For the Dutch case, which I have exploited in measuring the effect of capital constraints on venture performance, this assumption seems realistic since education is easily accessible for people with modest financial endowments too. My fifth assumption is that individuals, as well as policy makers, bankers and other parties involved, have no more insight in the returns to education and the effect of capital constraints than we as researchers have. This implies that individuals and policy makers share the knowledge (and common opinion) that the returns to education are similar for entrepreneurs and for employees and that capital constraints are binding.

The knowledge that the returns to education are high and that education is therefore a key input in a starting enterprise is informative for the design of educational policies and policies with respect to (selecting) starters designed by bankers and other capital suppliers. Moreover, the adequate design of tax and subsidy measures toward starters and their capital suppliers might also benefit from these insights.

Policy makers should be aware that the returns to education for entrepreneurs are higher than these of employees. Governments could take two actions regarding this new knowledge. They could invest in higher schooling for (prospective) entrepreneurs or they could invest in stimulating higher educated individuals to opt for entrepreneurship. The first action will make sure that entrepreneurs will perform better on average and that they will thereby generate more benefits which will not only benefit the entrepreneur but society as a whole. This will accordingly decrease the social costs pertaining to bankruptcy. The second action appeals to the fact that, at least in Europe, entrepreneurship seems not to be a favored option, or even part of the choice set, among young people with higher education. They usually favor working in a large multinational company and do not even think about self-employment. Therefore, governmental programs to stimulate the awareness of the option

of entrepreneurship to college and university students might be beneficial (see Chapter 4 of this volume).

Further research into the most beneficial types of education will lead to practically useful insights for the design of schooling tracks that are recommended to entrepreneurs. These types of education can then further be used to recruit entrepreneurs and to further stimulate and facilitate these entrepreneurs financially or otherwise, for instance, by means of subsidies and (loan) guarantees.

Improved policy measures can decrease entry barriers for potentially successful starters and also increase useful support to those starters. This will reduce the social costs of bankruptcy and increase the social benefits of innovative enterprises.

## NOTES

<sup>1</sup> Parts of this section overlap with Van der Sluis and Van Praag (2004).

<sup>2</sup> In signaling, the party with private information—that is, the employee in the selection and hiring process by employers—takes the lead in adopting behavior that, upon appropriate interpretation, reveals information about his productivity.

<sup>3</sup> See Van der Sluis, Van Praag and Vijverberg (2005) for a meta-analysis pertaining to less-developed countries.

<sup>4</sup> For instance Blanchflower and Oswald (1998); Burke et al. (2002); Cowling et al. (2004); Dunn and Holtz-Eakin (2000); Evans and Jovanovic (1989); Fairlie (1999); Gentry and Hubbard (2001); Henley (2004); Holtz-Eakin et al. (1994b); Hurst and Lusardi (2004); Johansson (2000); Lel and Udell (2002); Lindh and Ohlsson (1996); Moore (2004); Quadrini (1999); Uusitalo (2001); Van Praag and Van Ophem (1995).

<sup>5</sup> As in other studies (see Astebro and Bernhardt, 2003; Bates, 1990; Burke et al., 2002; Cooper et al., 1994; Cowling et al., 2004; Cressy, 1996; Evans and Jovanovic, 1989; Holtz-Eakin et al., 1994a; Hurst and Lusardi, 2004; Lindh and Ohlsson, 1996; Van Praag, 2003).

<sup>6</sup> The same categorization might be applied to research into the relationship between the probability of becoming an entrepreneur and access to capital. At least all approaches discussed here have been applied to the analysis of that research question.

<sup>7</sup> Empirical research into the relationship between assets and business start-up has recently taken an interesting turn, which I expect to be taken for studying the relationship between assets and performance too. Hurst and Lusardi (2004) and Moore (2004) found strong evidence that the relationship between assets and the probability of start-up is non-linear. In fact, assets are only a significant determinant of start-up for the highest percentiles of assets.

<sup>8</sup> However, HJR seem to have dealt with this issue in a neat way: by controlling for (i) whether the inheritance donor is an entrepreneur too and (ii) a measure of firm performance prior to the receipt of the inheritance.

<sup>9</sup> Though both participation in a lottery and selection into entrepreneurship are related to risk attitude and therefore to each other (see Cramer et al., 2002).

<sup>10</sup> Notably, there is one study, so far, that approaches the question of performance and liquidity constraints very differently: Astebro and Bernhardt (2003) estimate the effect of having a bank

loan on survival. They find that having a bank loan affects survival positively while controlling for other relevant factors.

<sup>11</sup> It is assumed that the positive effects of  $\theta_i$  on  $A_i$  and  $k_i^*$  just cancel out. Capital need ( $k_i^* - A_i$ ) is independent of  $\theta_i$  and does not affect  $\beta_i$  or  $P$ .

<sup>12</sup> They considered the 7% + 17% of the sample who answered yes as an alternative indicator of capital constraints. This weakened the result considerably. The same holds for the alternative specification where the first answer is translated into  $\beta_i = 1$ , the second into  $\beta_i = 0.5$  and the third is equivalent to  $\beta_i = 0$ .

<sup>13</sup> A third cause of a downward bias could be *overconfidence*. Overconfident entrepreneurs might aim at larger than efficient amounts of start-up capital. Without access to the desired amount, they feel constrained and report so. Unfortunately, BVD are unable to test this hypothesis that would again lead to an underestimate of the effect of the capital constraint on performance.

<sup>14</sup> The profit measure has zero as lower bound: Negative profits are not observed. Therefore, the equation is estimated using tobit regressions. For duration, BVD apply a log-logistic survival model.

## REFERENCES

- Ashenfelter, O. and Krueger, A. (1994). Estimates of the Returns to Schooling from a New Sample of Twins. *American Economic Review* 84 (5): 1157–1173.
- Ashenfelter, O., Harmon, C. and Oosterbeek, H. (1999). A Review of the Schooling/Earnings Relationship with Tests for Publication Bias. *Labor Economics* 6: 453–470.
- Astebro, T. and Bernhardt, I. (2003). Start-up Financing, Owner Characteristics, and Survival. *Journal of Economics and Business* 55: 303–319.
- Audretsch, D. and Keilbach, M. (2003). *Entrepreneurship Capital and Economic Performance*. Discussion Paper No. 3678, Centre for Economic Policy Research, London.
- Bates, T. (1990). Entrepreneur Human Capital Inputs and Small Business Longevity. *Review of Economics and Statistics* 72 (4): 551–559.
- Blackburn, M. and Neumark, D. (1993). Omitted-Ability Bias and the Increase in the Return to Schooling. *Journal of Labor Economics* 11 (3): 521–544.
- Blanchflower, D. and Oswald, A. (1998). What Makes an Entrepreneur? *Journal of Labor Economics* 16 (1): 26–60.
- Boadway, R., Marceau, N., Marchand, M. and Vigneault, M. (1998). Entrepreneurship, Asymmetric Information, and Unemployment. *International Tax and Public Finance* 5 (3): 307–327.
- Bonjour, D., Cherkas, L., Haskel, J., Hawkes, D. and Spector, T. (2003). Returns to Education: Evidence from U.K. Twins. *American Economic Review* 93 (5): 1799–1812.
- Burke, A., FitzRoy, F. and Nolan, M. (2002). Self-employment Wealth and Job Creation: The Roles of Gender, Non-pecuniary Motivation and Entrepreneurial Ability. *Small Business Economics* 19: 255–270.
- Cantillon, R. (1979). *Essai sur la Nature du Commerce en Général*. Kinokuniya, Tokyo.
- Cooper, A., Gimeno-Gascon, F. and Woo, C. (1994). Initial Human and Financial Capital as Predictors of New Venture Performance. *Journal of Business Venturing* 9: 371–395.
- Cosh, A., Cumming, D. and Hughes, A. (2005). *Outside Entrepreneurial Capital*. SSRN Working Paper.
- Cowling, M., Taylor, M. and Mitchell, P. (2004). Job Creators. *The Manchester School* 72 (5): 601–617.

- Cramer, J., Hartog, J., Jonker, N. and Van Praag, C. M. (2002). Low Risk Aversion Encourages the Choice for Entrepreneurship: An Empirical Test of a Truism. *Journal of Economic Behavior and Organization* 48: 29–36.
- Cressy, R. (1996). Are Business Startups Debt-Rationed? *Economic Journal* 106: 1253–1270.
- Davidsson, P. and Honig, B. (2003). The Role of Social and Human Capital Among Nascent Entrepreneurs. *Journal of Business Venturing* 18: 301–440.
- De Meza, D. and Webb, D. (2000). Does Credit Rationing Imply Insufficient Lending? *Journal of Public Economics* 78 (3): 215–234.
- Dunn, T. and Holtz-Eakin, D. (2000). Financial Capital, Human Capital, and the Transition to Self-employment: Evidence from Intergenerational Links. *Journal of Labor Economics* 18 (2): 282–305.
- Evans, D. and Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints. *Journal of Political Economy* 97 (4): 808–827.
- Evans, D. and Leighton, L. (1990). Small Business Formation by Unemployed and Employed Workers. *Small Business Economics* 2: 319–330.
- Fairlie, R. (1999). The Absence of African-American Owned Business: An Analysis of the Dynamics of Self-employment. *Journal of Labor Economics* 17: 80–108.
- Fazzari, S., Hubbard, R. and Petersen, B. (1988). Financing Constraints and Corporate Investment. *Brookings Papers on Economic Activity* 1: 141–195.
- Fredland, J. and Little, R. (1981). Self-employed Workers: Returns to Education and Training. *Economics of Education Review* 1 (3): 315–337.
- Fritsch, M., Brixey, U. and Falck, O. (2004). *The Effect of Industry, Region and Time on New Business Survival—A Multi-Dimensional Analysis*. Working Paper 4-2004, Freiberg University, Germany.
- Gentry, W. and Hubbard, R. (2001). *Tax Policy and Entrepreneurial Entry*. Working Paper, Columbia University.
- Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica* 45 (1): 1–22.
- Hamilton, B. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-employment. *Journal of Political Economy* 108 (3): 604–631.
- Henley, A. (2004). Self-employment Status: The Role of State Dependence and Initial Circumstances. *Small Business Economics* 22: 67–82.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. (1994a). Sticking it Out: Entrepreneurial Survival and Liquidity Constraints. *Journal of Political Economy* 102 (1): 53–75.
- Holtz-Eakin, D., Joulfaian, D. and Rosen, H. (1994b). Entrepreneurial Decisions and Liquidity Constraints. *Rand Journal of Economics* 25 (2): 334–347.
- Hurst, E. and Lusardi, A. (2004). Liquidity Constraints, Household Wealth and Entrepreneurship. *Journal of Political Economy* 112 (2): 319–347.
- Johansson, E. (2000). Self-employment and Liquidity Constraints: Evidence from Finland. *Scandinavian Journal of Economics* 102 (1): 123–134.
- Kirzner, I. (1973). *Competition and Entrepreneurship*. University of Chicago Press, Chicago.
- Knight, F. (1971). *Risk, Uncertainty and Profit*. University of Chicago Press, Chicago.
- Lazear, E. (2004). Balanced Skills and Entrepreneurship. *American Economic Review, Papers and Proceedings* 94: 208–211.
- Le, A. (1999). Empirical Studies of Self-employment. *Journal of Economic Surveys* 13 (4): 381–416.
- Lel, U. and Udell, G. (2002). *Financial Constraints, Start-up Firms and Personal Commitments*. Mimeo, Kelly School of Business, Indiana University.
- Leroy, S. and Singell, L. (1987). Knight on Risk and Uncertainty. *Journal of Political Economy* 95: 394–406.

- Leuven, E., Oosterbeek, H. and van der Klaauw, B. (2003). *The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment*. Discussion Paper 38/03, Tinbergen Institute, The Netherlands.
- Lindh, T. and Ohlsson, H. (1996). Self-employment and Windfall Gains: Evidence from the Swedish Lottery. *Economic Journal* 106: 1515–1526.
- Marshall, A. (1930). *Principles of Economics*. Macmillan, London.
- Meyer, J. and Kuh, E. (1957). *The Investment Decision: An Empirical Study*. Harvard University Press, Boston, MA.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. Columbia University Press, New York.
- Modigliani, F. and Miller, M. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review* 48: 261–297.
- Moore, K. (2004). *Do Liquidity Constraints Matter for New Entrepreneurs?* Discussion Paper, Federal Reserve Board Finance and Economics.
- Myers, S. and Majluf, B. (1984). Corporate Financing and Investment Decisions when Firms Have Information that Investors do not Have. *Journal of Financial Economics* 13: 187–221.
- Oosterbeek, H. (1992). *Essays on Human Capital Theory*. Thesis Publishers, Amsterdam.
- Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*, Cambridge University Press, Cambridge.
- Parker, S. C. and Van Praag, C. M. (2006). Schooling, Capital Constraints and Entrepreneurial Performance: The Endogenous Triangle. *Journal of Business and Economic Statistics*, forthcoming.
- Quadrini, V. (1999). The Importance of Entrepreneurship for Wealth Concentration and Mobility. *Review of Income and Wealth* 45: 1–19.
- Riley, J. (2002). Weak and Strong Signals. *Scandinavian Journal of Economics* 104: 213–236.
- Robinson, P. and Sexton, E. (1994). The Effect of Education and Experience on Self-employment Success. *Journal of Business Venturing* 9: 141–156.
- Say, J.-B. (1971). *A Treatise on Political Economy or the Production, Distribution and Consumption of Wealth*. A. M. Kelley Publishers, New York.
- Scherr, F., Sugrue, T. and Ward, J. (1993). Financing the Small Firm Startup: Determinants of Debt Use. *Journal of Small Business Finance* 3 (1): 17–36.
- Schumpeter, J. (1934). *The Theory of Economic Development*. Harvard University Press, Cambridge MA.
- Silva, O. (2004). *Entrepreneurship: Can the Jack-of-all-Trades Attitude be Acquired?* Discussion Paper No. 665, Centre for Economic Policy, London.
- Spence, A. (1973). *Market Signaling: Information Transfer in Hiring and Related Processes*. Harvard University Press, Cambridge, MA.
- Taylor, M. (1999). Survival of the Fittest? An Analysis of Self-employment Duration in Britain. *Economic Journal* 109: 140–155.
- Tucker, I. (1987). The Impact of Consumer Credentialism on Employee and Entrepreneur Returns to Higher Education. *Economics of Education Review* 6 (1): 35–40.
- Tucker, I. (1985). Use of the Decomposition Technique to Test the Educational Screening Hypothesis. *Economics of Education Review* 4 (4): 321–326.
- Uusitalo, R. (2001). Homo Entrepreneurus? *Applied Economics* 33: 1631–1638.
- Van der Sluis, J. and Van Praag, C. M. (2004). Economic Returns to Education for Entrepreneurs: The Development of a Neglected Child in the Family of Economics of Education? *Swedish Economic Policy Review* 11 (2): 183–226.
- Van der Sluis, J., Van Praag, C. M. and Vijverberg, W. (2005). Entrepreneurship Selection and Performance: A Meta-analysis of the Impact of Education in Less Developed Countries. Forthcoming in the *World Bank Economic Review*.

- Van der Sluis, J., Van Praag, C. M. and Van Witteloostuijn, A. (2004). *The Returns to Education: A Comparative Study between Entrepreneurs and Employees*. Working Paper, University of Amsterdam.
- Van der Sluis, J., Van Praag, C. M. and Vijverberg, W. (2003). *Entrepreneurship Selection and Performance: A Meta-analysis of the Impact of Education in Industrialized Countries*. Discussion Paper, Tinbergen Institute, University of Amsterdam.
- Van Praag, C. M. (1999). Some Classic Views on Entrepreneurship. *De Economist* 147 (3): 311–335.
- Van Praag, C. M. (2003). Business Survival and Success of Young Small Business Owners: An Empirical Analysis. *Small Business Economics* 21 (1): 1–17.
- Van Praag, C. M. and Van Ophem, J. (1995). Determinants of Willingness and Opportunity to Start as an Entrepreneur. *Kyklos International Review for Social Sciences* 48: 513–540.
- Van Praag, C. M., De Wit, G. and Bosma, N. (2005). Initial Capital Constraints Hinder Entrepreneurial Venture Performance: An Empirical Analysis. *Journal of Private Equity* 9 (1): 36–44.
- Wagner, J. (2002). *Testing Lazear's Jack-of-all-Trades View of Entrepreneurship with German Data*. Discussion Paper No. 592, IZA, Bonn.
- Weick, K. (1996). Drop your Tools: An Allegory for Organizational Studies. *Administrative Science Quarterly* 41: 301–314.
- Wolpin, K. (1977). Education and Screening. *American Economic Review* 67: 949–958.

MAKS TAJNIKAR, PETRA DOŠENVIČ BONČA and LIDIJA ZAJEC

*Faculty of Economics, University of Ljubljana, Kardeljeva ploščad 17, 1109 Ljubljana, Slovenia*

*E-mails: maks.tajnikar@ef.uni-lj.si; petra.d.bonca@ef.uni-lj.si; lidija.zajec@ef.uni-lj.si*

## 18. Harvesting in High Growth Firms

### 1. INTRODUCTION

The term harvesting is generally associated with entrepreneurship and entrepreneurship research. Though many definitions of harvesting can be found in the literature, all of them state it is always an activity through which an investor draws a return on capital from a specific investment. Authors often link harvesting to terms such as a divestiture, withdrawal or exit from an investment, although it will on occasion be defined in a wider sense and is not always related to an exit. It is therefore possible to systematize a variety of forms of harvesting and not all of them relate to investors withdrawing from a business. Conceptual work on harvesting began with sporadic publications that came out of Babson College. Timmons (1990, 1999), Bygrave and Timmons (1992), Petty (1994), Petty et al. (1994), MacIntosh (1997) and Cumming and MacIntosh (2001, 2003a, 2003b) provided a primary theoretical framework specific to exits of venture capital. Their work is chiefly based on Black and Gilson (1998) and Gompers and Lerner (1999a, 1999b, 2004). Extending this theory to a more generalized perspective of any type of investor, the question that is nowadays almost certainly one of the most interesting in the harvesting area is one that asks what determines the investor's choice of a particular form of harvesting or, in other words, what are the determinants of the choice of a particular form of harvesting. This is also the topic of this chapter.

At first glance, it would be logical to expect the choice of harvesting form to be subject to the will and interests of the investor; it is, after all, an issue of their returns. Although the selection of a specific form of harvesting should principally be a reflection of the personal interests of the investors and owners, it can and should be assumed that certain characteristics of the investor and the circumstances influencing their decision-making process have some



bearing on these personal interests and, as a result, on the harvesting form they choose. In addition, the characteristics of both the firm and its environment determine the potential and limitations for which the interests and expectations of the owners and investors can be satisfied. The firm's internal and external characteristics also determine its growth opportunities which indisputably affect the possibility of realizing investors' interests. Therefore, the form of harvesting chosen emerges from the relationship between the interests of the owners and investors on one hand, and the firm's internal and external characteristics on the other. This chapter aims to identify the key factors that determine the choice of a particular harvesting form. These factors are collected into four broad groups; it will be claimed that these four groups can determine the selection of any of the five forms of harvesting that will be discussed in the body of this chapter.

Three sections follow this one. Section 2 defines and discusses the meaning and application of the term harvesting. Section 3 presents a literature review aimed at exploring all the various forms of harvesting and giving a structured overview of those factors that affect the choice of a particular harvesting form. The final section concludes.

## 2. DEFINITION AND GENERAL CHARACTERISTICS OF HARVESTING

In the entrepreneurship literature, the term harvest was initially used as a synonym for the term exit from an investment (Petty et al., 1994). Petty (1994) also defined a harvest "as the owners' and investors' strategy for achieving the terminal after-tax cash flows on their investment." From this viewpoint, harvesting is the final phase of the investment process. According to Petty et al. (1999), the latter is a three-step process of creating wealth in which the first step is building the business, the second is making it grow and harvesting is the third. Similarly, Isaksson (1998) observed these as phases of the venture capital process and referred to them as the phase of investing, the phase of value-adding and the phase of exiting.

Harvesting can also be regarded as more than just selling and leaving the business (Petty et al., 1999). King (2002) defined it as the "path to realizing the gains from an investment." This implies that harvesting may involve activities other than exiting an investment. It can, in more general terms, be considered the activity by which investors draw their profit from a particular investment with the intention of using the profit for other purposes. An investor may thus harvest the returns of their investment through either exiting their investments or by gradually extracting the firm's free cash flows over time.

In order to be successful in reaping the rewards of the investment process, harvesting has to be a carefully planned activity. It requires the

definition of a clear harvesting strategy which is a plan for how an investor will realize a return on their investment, namely, at the time of the harvest, a material amount of the economic value created over the life of the firm can be either preserved or lost (Petty et al., 1994). That is why designing the harvesting strategy is necessary if the firm's goal is not only to provide a living for the entrepreneur but to both create and extract value from a growing and profitable business for its founder and other investors (Petty, 1994). The importance of developing harvesting strategies largely derives from the conviction that a firm is successful if its results, seen from economic, social and personal points of view, satisfy or even exceed the expectations of all those individuals who have in one way or another expressed their interest in it (Ekins et al., 1992; Gianaris, 1996; Hooper and Potter, 1997; Van Osnabrugge and Robinson, 2000; Wickham, 1998). Their expectations cannot be met if the created value remains locked and unharvested. The need to develop a harvesting strategy is also well reflected by the evidence that suggests venture capitalists have been better at investing than harvesting their investments. For example, a survey by Wall and Smith (1998) compared the level of exits with the level of new investments for venture-backed firms in Europe to find that the amount of money being invested in portfolio firms by European venture capital funds far exceeds the amounts harvested and, as they show, many European investors who planned their exit from day one claimed a high success rate in satisfying the expectations of their stakeholders, while most of those that gave little thought to defining a specific exit route in advance did not.

During harvesting, the "harvest potential" of businesses, which is "the ability to reap the rewards of the investment process," is fully revealed (Sahlman et al., 1999). Through harvesting, the ultimate value is created for all participants in the venture, especially the owners, managers and employees (Petty, 1994). In this context, harvesting can be seen as creating the value and other benefits for all stakeholders directly involved with the venture. It can, for example, enable venture capital fund investors to evaluate the quality of their venture capitalists' reputation capital (Megginson and Weiss, 1991). Also, through harvesting "the seeds of renewal and reinvestment are sown" (Timmons, 1999). This implies that harvesting can also be regarded as the activity of reinitiating the entrepreneurial process or, more precisely, according to Timmons (1999), as the "recycling of entrepreneurial talent and capital" or, according to Black and Gilson (1998), as the recycling of also a venture capitalist's non-financial contributions from successful companies to early stage companies.

The above conclusions imply that developing a harvesting strategy is important for the very different individuals who are the firm's stakeholders. These are the firm's owners, managers, employees, suppliers, customers, local authorities, etc. Extensive literature also highlights the role of venture capital

fund investors as important stakeholders. Yet professional venture capitalists are not the only investors in new businesses. Other individuals such as family members and other private informal investors often also assume this role (Prasad et al., 1995). The timing and selected form of harvesting impact on the firm's different stakeholders in different ways and that is why they attempt to influence the harvesting strategies that are pursued.

Even though the firm has many stakeholders who express their interest in the firm, the prevailing interest is indisputably that of the firm's investors and owners (Petty, 1994; Timmons, 1999). This is because their role is crucial for the firm's creation since through the markets of production factors with their investment they enable the integration of all the elements making up the firm. A firm must therefore basically act so that it realizes the interests of the investors who back a business. These can either be friends, family, business angels, institutions or venture capitalists (King, 2002). From the investors' point of view, a firm is thus founded and exists in order to realize a harvest that leads to the maximum possible increase in their assets which is the main interest of any rational investor. Petty et al. (1999) supported this view by emphasizing that a firm's owners and investors will be denied a significant amount of its value without the opportunity to harvest. This implies that the harvesting strategy is not a strategy initially formed at the firm level. It is, in fact, the strategy of the business founders and investors aimed at gaining liquidity from investments of money and effort in order to meet the need for the business to grow, the consumption requirements of its founders and other investors, or the challenges of tax and estate planning (Kensinger et al., 2000). The firm's strategy, however, has to incorporate the harvesting strategy of the business founders and investors if it is to satisfy the expectations of its key stakeholders.

According to Petty (1994), a well-conceived harvesting strategy of the business founders and investors consists of three elements which are the answers to three key questions. These questions are: why to harvest; when to harvest and in what form? Giot and Schwienbacher (2003) also mentioned the elements of a harvesting strategy but they only referred to its timing and form. Defining the harvesting strategy and all its elements, that is, reason, time and form, according to Timmons (1999), calls for patience, a vision, realistic valuation and outside advice.

Considering that the harvesting strategy is a strategy of the owners and investors, they give the answer to the question of why to harvest. Namely, it is through harvesting that they can obtain the resources needed for various purposes determined either subjectively or objectively. Investors may thus decide to harvest because they believe the firm is at its best harvest potential. Their decision to harvest, however, can also be motivated by other factors. Kensinger et al. (2000), for example, highlighted that, in cases where the owners borrow against future cash flows of their investment, the limits on the time

horizon of their borrowing are likely to affect their ability and demands to realize the full liquidity potential of the business. Another example here is given by Sahlman and Hurlock (1992) who showed that this element of a harvesting strategy is also influenced by the investors' status which relates to their personal wealth, indebtedness and other characteristics.

The owners and investors also give the answer to the question of when to harvest. According to Sahlman (1988), the assumed horizon date for the harvest should be an integral part of the financial contract between the business founders and investors. Either formally or informally the investors should let the entrepreneur or management know that they will want to harvest their investment at a certain time in the future when their expectations are met. However, the time to harvest cannot only be determined by the preferences of the owners and investors. The investment duration should also be determined so that when the time for harvesting arrives, information asymmetries between the sellers and the new owners are minimized thereby maximizing the reward of the existing owners and investors (Cumming and MacIntosh, 2003c). It also requires the investors' understanding and ability to take advantage of a so-called strategic window (Timmons, 1999). This refers to the time when the most suitable circumstances emerge for the actual harvest to be realized. It is believed that it takes about seven to ten years for a strategic window to open and it is in this period that investors can realize the most attractive returns on their investment (Petty, 1994). For venture capital investments Sahlman (1990) showed that it takes approximately five years before investments are mature enough to be harvested. In the European case, as noted by Wall and Smith (1998), venture capitalists believe a normal target life for their investment is between 3 and 6 years, although in reality the existing venture capital portfolio would take eight years to divest. Mason and Harrison (2002) showed that for the U.K. case the median time to exit for successful investments is four years. It is also important to note that the window of opportunity for harvesting can open and close quickly (Petty et al., 1999). The timing of a strategic window is determined by several factors. Cumming and MacIntosh (2001), for example, showed that U.S. firms with a venture capital investment in the early stage of development have a shorter average investment duration and their U.S. data also indicated that the greater availability of capital to the venture capital industry resulted in a shorter average investment duration. Later research by Cumming and MacIntosh (2003c) confirmed their previous findings and added that issues such as whether the exit was preplanned, whether the exit was made in response to an unsolicited offer and certain institutional factors also determine investment duration. Das et al. (2003) also highlighted the effect of the stage of development on the investment's duration as they showed that, for over two-thirds of late-stage companies, a successful exit happens within three years of financing and only one-third of early-stage companies had a liquidity

event within three years of financing. Giot and Schwiendacher (2003) further showed that the time to exit also depends on the type of industry. They found that Internet firms had the fastest IPO exits and were also the fastest to liquidate in case they were unsuccessful. They further determined that biotech firms are also characterized by fast IPO exits but are slow to liquidate.

Harvesting can emerge in very different forms. These include free cash flow, dividend or current profit payouts, an initial public offering, trade sale, management buyout, employee buyout, management buy-in, acquisition, merger, takeover and liquidation of the company or sale of its assets. Black and Gilson (1998), Schwiendacher (2005), Fleming (2002), Cumming and MacIntosh (2003a, 2003b) rank ordered the exit vehicles according to the quality of the entrepreneurial firm. They list IPOs first, followed by acquisitions, secondary sales, buybacks and write-offs. According to Pagano et al. (1998), one of the many driving principles underlying this ordering relates to the ability of harvesting forms to resolve information asymmetries. Different forms of harvesting occur with different frequencies and probabilities. Wall and Smith (1998), for example, examined venture capital exits in 1994 and 1995 and found that by far the largest number of exits took place in the form of a trade sale. Das et al. (2003) studied private U.S. firms in the period between 1980 and 2000 and found that the probability of an exit via an IPO is roughly 20–25% for firms financed in an early stage, expansion stage or later stage. They determined that the probability of an exit via an acquisition is approximately 10–20% and that this probability is much higher for firms in later stages. In addition, they showed that 44% of companies in late-stage financing experienced a liquidity event and that only 34% of early-stage firms had a successful exit. Selecting different harvesting forms is also conditional on the type of investor involved. Venture capital investments typically do not pay dividends; rather, returns are derived from capital gains upon exit (Cumming et al., 2004). Venture capitalists prefer to take their investments public as 30% of firms backed by venture capitalists over the past two decades have gone public (Gompers and Lerner, 2004). Mason and Harrison (2002), on the other hand, showed that IPOs are a less common form of harvesting for business angels. Erikson and Sørheim (2005) supported this finding by showing that technology angels prefer to exit through trade sales.

Investors also differ in their approaches to developing their harvesting strategies and their above-mentioned elements. Some do not like to think about the harvest, while others begin the venture with the goal of harvesting it (Petty et al., 1999). Therefore, some investors are motivated exclusively by the prospect of successfully harvesting an investment while others either give little thought to harvesting or may not be in a position to make a decision regarding an exit. Wall and Smith (1998) referred to these two types of investors as the proactive and the passive investor, respectively. King (2002), for example, noted that entrepreneurs do not give much thought to the eventual outcome in the case

of family businesses expected to run for successive generations. Kaplan (2003) also pointed out that entrepreneurs start thinking about an exit plan when the stress of managing the business becomes too great or when they lack enough time for the firm due to family commitments. King (2002) further illustrated that there are also many reactive investors who are happy to “sit back as long as the interest or dividends keep coming in.” Also, business angels can be “driven as much by their desire to apply their know-how to a fledgling firm as they are to think about realizing an exit” (King, 2002). Owner-managers in a private firm, however, may be unable to make a decision to sell as other shareholders might be unwilling or unable to buy. Holmburg (1991) and Hyatt (1990), on the other hand, found evidence that harvesting strategies are considered either formally or informally in most cases. The investors’ approach to developing a harvesting strategy is revealed by their exit behavior that shows how the investors handle the exit during the investment process. Relander et al. (1994) identified two patterns of exit behavior: the “path sketcher” and the “opportunist.” The first is working actively with exit problems during the whole investment process while the other “trusts that their management skills and the concept of the investment target will lead to an exit opportunity” at just the right time. In the latter case, a decision to harvest can also result from an unexpected crisis (Petty et al., 1999). Some investors, therefore, merely react to created harvesting opportunities while others carefully plan their harvesting strategy. Case studies of King (2002) showed that, in practice, the exit generally occurs as a combination of astute planning and opportunism.

In the literature, harvesting is not considered to be part of just any investment process. It is seen as an element of the entrepreneurial investment process (Smith and Smith, 2000), the investment process in which an important role is also assumed by the entrepreneur. As already shown, harvesting is particularly important to outside investors. The ability to make a profitable exit is the main motive for outside investors (Sahlman, 1990; Gompers and Lerner, 2004). For entrepreneurs, however, harvesting has both financial and personal (nonfinancial) considerations and a harvesting strategy is thus essential to both the entrepreneur’s personal and financial success (Petty et al., 1999). That is why the harvest goals of outside investors and the entrepreneur may differ. For the entrepreneur–investor, the rate of return is still the central issue but the personal effects on the entrepreneur’s life are also of great importance. “Harvesting the venture brings both excitement and trepidation to most entrepreneurs” (Petty, 1994). The outside investor’s exit is not always also the entrepreneur’s withdrawal from the portfolio firm (Isaksson, 1998). From this viewpoint, forming a harvesting strategy is also essential for the achievement of the entrepreneur’s personal success (Petty et al., 1999). Even in circumstances where the entrepreneur is not simultaneously the investor, harvesting can have

a great impact on the entrepreneur's life considering that it can change or even terminate his position.

If the common characteristic of entrepreneurs and outside investors is that they all own equity, then this gives them a strong incentive to maximize the firm's value. Both the entrepreneur or manager and investor thus share the risks of the firm and want to make a return on their investment without jeopardizing the business' future. Yet their priorities differ. The first is focused on creating value and the other on realizing it (King, 2002). That is why they may have incentives to pursue different harvesting strategies. Berglöf (1994), for example, noted that the choice among different forms of harvesting may have important distributional consequences for the entrepreneur and the venture capitalist, as different harvesting forms are accompanied by different sets of cash flows. For venture capitalists, an exit is not optional. Venture capital funds have a fixed life, usually of ten years with an option to extend for a period up to three years (Gompers, 1996). Entrepreneurs, however, receive benefits such as freedom and security from operating a privately held firm that are not available to the firm's investors and may thus be perfectly happy to preserve the status quo. Berglöf (1994) thus suggested that such potential conflicts between investors and entrepreneurs may be mitigated by allocating the decision to sell the firm or the initiation and the veto right to the most vulnerable party. Smith (2001), for example, points out that often venture capitalists allow entrepreneurs to prove the business concept but with the passage of time venture capitalists begin to exert more control over harvest decisions and can select a form of harvesting that is not necessarily also preferred by the entrepreneur. Considering that entrepreneurs may not want to stop their projects voluntarily and have a preference for continuation, venture capitalists can take the right to decide on the form and the timing of harvesting themselves (Schwienbacher, 2005). This potential conflict between investors and entrepreneurs highlights the importance of designing a harvesting strategy as it can be a means for the entrepreneurs or managers and investors to align their interests, develop a shared view of the investment horizon and the way value is to be realized (King, 2002).

It is precisely because of this personal involvement of the entrepreneur, and not only the outside investor, that harvesting and harvesting strategies are mostly debated in the literature on entrepreneurship. Given that the harvesting strategy is not a strategy formed at the firm level but at the level of individuals—owners and investors—it is usually not dealt with in the strategic management of firms, which is a technique used for creating the firm's desired future (Didsbury, 1996). Strategic management is the creation, realization and evaluation of functional decision making which enables organizations to achieve their set aims (David, 1999). It thus consists of a series of decisions and actions that provide the conditions for forming and realizing plans for the achievement of business aims (Pearce and Robinson, 1991; Sharplin, 1985). The fact that

strategic management does not deal with the formation of harvesting strategies is clearly shown by the fact that the issue of harvesting is basically not covered in the strategic management literature.

Nevertheless, even though the topic of harvesting strategies is debated in the entrepreneurship literature and is a strategy formed not at the firm level but at the individual level, it is still closely linked to strategic management. As already shown, harvesting strategies should underpin strategic management if the latter is to satisfy the expectations of its key stakeholders. Strategic management creates the conditions in which firms direct their energy and resources to the specification and realization of business aims that derive from those interests of owners and investors that are primarily reflected in their harvesting strategy. The ultimate goal of corporate governance is to ensure that investors receive a return on their investment (Shleifer and Vishny, 1997). A firm is thus founded and exists in order to harvest the owners' investment. The processes of building and making a business grow, which are in the focus of strategic management, therefore cannot exist without owners and investors and their harvesting strategy. That is why these processes have to be in line with the owners' and investors' interests, strategic management has to take the harvesting strategy of the owners and investors into consideration and the firm's strategy must take the individual strategies of owners and investors into account.

As previously indicated, the owners and investors form a harvesting strategy with the aim of realizing the harvest that leads to the maximum possible increase in their assets. In order for the investment in the firm to bring the maximum possible increase in the investor's assets, the firm's primary goal must be its growth (Pearce and Robinson, 1991; Napuk, 1993). For entrepreneurs, it can even be stated that they "get their kicks from growing the firm; they know the payoff will take care of itself if they concentrate on the money-making part of the process" (Timmons, 1999). Namely, business growth on one hand enables the growth of the firm's assets and thus also the nominal value of the firm's equity that represents the owners' assets. On the other hand, business growth has an impact on the market value of the firm's equity and thus creates the conditions that enable the investor to harvest the most attractive returns on their investment.

It is, however, important to note that the extent to which growth possibilities are recognized is primarily determined by the firm's management. A specific form of harvesting, therefore, emerges not only as an expression of the will of investors but is also a result of the managers' capacities to pursue the firm's growth possibilities. Some findings show that the latter are determined by managerial characteristics such as years of experience in marketing/sales, willingness to take risks and tolerance of ambiguity (Pearce and Robinson, 1991). In order to explain the choice of a particular harvesting strategy it is therefore necessary to link the interest of the owners with both the capacities of



the managers and their interests (Braganza and Lambert, 2000). This opens up an additional set of questions regarding active and passive firm growth strategies (Morris and Sexton, 1996; Tajnikar and Brščič, 2002). The first occur when managers are capable of creating suitable conditions for growth within and outside the firm, and we talk of the second when managers are only capable of exploiting conditions within and outside the firm that emerge independently of them. The more managers are able and willing to exploit conditions for the firm's passive growth, or the maximum extent to which managers are able and willing to ensure the firm's growth by forming an active growth strategy, the greater is the possibility that investors achieve the harvesting forms that are most in accordance with their planned and expressed interests.

The choice of a particular harvesting strategy cannot only be explained by linking the interest of the owners with the capacities of the managers to pursue the firm's growth possibilities given that the latter are also determined by the firm's internal and external characteristics. This thus clearly shows that the firm's internal and external characteristics which determine the firm's growth possibilities also have to be taken into consideration. Research supports this conclusion as it confirms that the likelihood of different forms of harvesting depends not only on firm-specific characteristics but also on the firm's environment. A particular form of harvesting, therefore, emerges in the relationship between the interests of the investors on one hand, and those internal and external characteristics of the firm which determine the firm's growth possibilities on the other. These characteristics of the firm and its environment are either created by the firm's management or the firms must accept them as given circumstances of their business.

### 3. HARVESTING FORMS AND DETERMINANTS OF THEIR CHOICE

There are a number of ways to gather returns from a business and they are represented by different forms of harvesting. The breadth of the term harvesting is of great consequence when we proceed to find all possible forms of harvesting and seek to find the determinants of choice between them. It would be wrong to look upon harvesting as an activity that only comes into place upon the extinction or termination of a business opportunity. The process of designing harvesting strategies is frequently derived from the conviction that a business that is successful today cannot last forever. Hence harvesting is often understood as the investors' and owners' withdrawal from the business (Petty, 1994; Smith and Smith, 2000; Timmons, 1999). This understanding of harvesting is illustrated by Pearce and Robinson's (1991) distinction between the opposing "invest to grow" strategy and a "harvest" strategy, or by Thompson and Strickland's (1998) singling out of "harvest-divest" strategies, compared

to three other types of “invest-and-grow,” “fortify-and-defend” and “overhaul-and-reposition” strategies. These taxonomies all support the dichotomy between the intention to stay with an investment and the intention to withdraw from an investment thereby divesting or exiting a business. It is, however, important to note that the above-mentioned strategy classifications are found in the strategic management literature. This implies that they are the strategies of managers and not of the firm’s owners and investors. That is why, from the stance of investors and owners, defining a harvest as a synonym for exit is insufficient. We can thus also include in the whole array of harvesting forms those forms of harvesting that do not entail the withdrawal of investors from the business. Harvesting, therefore, does not necessarily mean that the entrepreneur or investors will leave the firm; the selected harvesting form merely defines how they will extract some or all of the cash flows from the investment to be used for other purposes (Petty, 1994). Hence, in the broader sense of the term, we can consider among harvesting forms all mechanisms used by investors to draw profit from a particular investment.

Various authors define and provide different taxonomies of harvesting forms. Timmons (1990), for example, distinguished “six principal avenues by which a firm can realize a harvest from the value it has created.” He considered the capital cow, the employee stock ownership plan, the management buyout, the merger, the outright sale and the public offering as six harvest options. In a similar way, Petty (1994) considered extracting free cash flows, a management buyout, an employee stock ownership plan, merging or being acquired and an initial public offering as forms of harvesting. MacIntosh (1997) outlined five different ways to exit a venture capital investment: an initial public offering, a trade sale or an acquisition, a secondary sale, a buyback or a management buyout and the final group which includes write-offs, reconstructions, bankruptcies and liquidations. MacIntosh (1997) and Cumming and MacIntosh (2003a) further differentiated between full and partial exits through initial public offerings, acquisitions, buybacks, secondary sales and write-offs when referring to the ways from which a venture capital investment can be exited. Cumming et al. (2004) also listed these five harvesting forms, but they added to the discussion of harvesting forms by dividing them into public exits (initial public offering) and private exits (all other harvesting forms). Petty et al. (1999) listed three groups of harvesting forms. These include selling the firm outright, systematically withdrawing the firm’s cash flows and releasing them to its owners, and either an initial public offering or a private placement of stock. They did, however, divide the first group (i.e. selling the firm outright) into strategic sales, financial sales and employee buyouts. In their classification, bust-up leveraged buyouts, build-up leveraged buyouts and management buyouts are considered as types of financial sales. Kensinger et al. (2000) included orderly liquidation, debt issue (preventing the cash flows of a mature business from

being reinvested in negative NPV projects), private sale to another firm, group of investors, managers, employees or family members, strategic or nonstrategic acquisition and initial public offering in their classification of harvesting forms. According to Smith (2001), most venture capitalists exit from investments either by selling shares pursuant to an acquisition of the portfolio firm, by selling or distributing shares after the portfolio firm completes an initial public offering, by the redemption of a venture capitalist's shares pursuant to a contractual "put" right or through liquidation of the portfolio firm and concomitant distribution of cash. Similarly, Gladstone and Gladstone (2004) listed six basic approaches to the cashing out of an investment. These include going public, selling to a strategic or financial buyer, selling back to the firm, selling to another investor, reorganizing the firm and liquidating the firm.

If we take the above-mentioned classifications of harvesting forms into consideration and include both those forms that come into play when investors exit the business, as well as those that do not assume the investors' withdrawal, we can enumerate five groups of harvesting forms. First, harvesting not associated with an exit from a business is in the form of free cash flow and dividend payouts to investors. Second, forms of harvesting through public offerings of stock encompass both public offerings (POs) and initial public offerings (IPOs). Third, the most diverse group of forms of harvesting includes a trade sale, a buyback, a management buyout (MBO), an employee buyout and a management buy-in (MBI). Fourth, the next group of harvesting forms includes mergers, acquisitions and takeovers. Fifth, the last group of harvesting forms where the business as a going concern ceases to exist, typically consists of bankruptcy, a firm's liquidation, sale of the firm's assets and a write-off.

The determinants that characterize the choice of a particular form of harvesting have been studied by several authors. The early works of Timmons (1990) and Petty (1994) were followed by MacIntosh (1997) and Cumming and MacIntosh (2001, 2003a, 2003b) who made attempts to provide a general theory of venture capital exits. The work of these two authors is primarily based on Black and Gilson's (1998) and Gompers and Lerner's (1999a, 1999b, 2004) contributions. MacIntosh (1997) in this manner analyzed factors that affect the choice of all types of venture capitalists' exits. This work has been extended by Cumming and MacIntosh (2001, 2003a, 2003b) using data for Canada and the U.S. For Europe, factors of investment exits were analyzed by Cumming (2002) and Scwienbacher (2005). For Australia, this contribution was made by Fleming (2002) and for the U.S. by Petty et al. (1999). These authors thus made attempts to integrate the full range of harvesting forms into their work, yet they were limited by their data. So far, most empirical work has been limited to the determinants of one or a few harvesting forms.

Studies of harvesting forms are under the influence of country specifics. Namely, in some cases studies reveal opposing results regarding the impact of a

certain harvesting form's determinants and they also show that the importance ascribed to different determinants can differ between countries. King's work (2002), for example, pointed to the fact that the investing of private equity in emerging markets differs from that in the U.S. and Europe and it is precisely the harvesting phase that is an entirely different and more complex process. Harvesting behavior differs significantly not only in countries in transition but also when comparing American and European economies as shown, for example, by Wall and Smith (1998), Cumming (2002), Botazzi and Da Rin (2003) and Schwienbacher (2005). Many participants in the European venture capital industry believe that much of the investment-divestment imbalance is caused by the relative scarcity of viable harvest options in Europe (Petty et al., 1999).

### *3.1. Free Cash Flow and Dividend Payouts*

The least frequently considered group of harvesting forms is that which includes free cash flow and dividend payouts to investors. They are those types of harvesting forms by which investors retain their positions in the firm. Certainly investors may choose not to realize the returns on their investment in a one-time payout or in the form of a capital gain. They may simply choose to collect gains through annual payouts on the equity they possess. For those payouts to be possible, the firm must be able to generate a stream of free cash flows over a prolonged time period. Timmons (1999) for this reason referred to those firms as cash cows. Payments to investors usually take the form of dividend payouts. Owners who harvest in such a way, therefore, require the practice of the long-term maximization of the free cash flows or dividend payouts of their firms.

Authors analyzing the factors of the choice of this group of harvesting forms find these factors in particular favorable circumstances of the firms that applied such strategies. These factors stem either from the environment, are attributable to the characteristics of the firm or the mode and intensity of their growth. Certain circumstances may also prohibit the choice of any other harvesting forms—those that would require a singular act of an investor exit and a capital gain payout.

Petty (1994) was one of the first to draw attention to the important fact that the selection of this group of harvesting forms might simply be the result of the inability to find an appropriate buyer. Sometimes there may be no buyers in the market, at other times the buyers may be unsuitable. In the former case, firm size also plays an important role (Timmons, 1999) as the high value of a firm which may also be a result of its size or the amount of equity engaged may be prohibitive. Lerner (1994) nevertheless pointed to the fact that circumstances characterized by the absence of buyers are actually situations in which potential

buyers are unwilling to pay the amounts expected by the present investors. He concluded that investors are likely to refrain from using equity markets to finance growth when equity values are lower. Another advantage of harvesting in the form of dividends is that no money or energy are expended in seeking out a buyer (Petty et al., 1999).

From the perspective of firms' characteristics, Petty (1994) and Petty et al. (1999) claimed that investors also choose this group of harvesting forms if they intend to retain control over the firm while they harvest their investment. They particularly had investors-entrepreneurs in mind since they contended that this harvesting form demands an entrepreneur's patience and willingness to harvest free cash flows over time (Petty, 1994, Petty et al., 1999). Cumming and MacIntosh (2003a) believed that conditions in favor of this group of harvesting forms come into play after a firm's IPO. The new investor (for instance, a venture capitalist) may, following the IPO, dispose of an investment by making dividend payouts on the investee firm's shares to the founding owners. Otherwise, venture capital investments typically do not pay dividends; rather, returns are derived from capital gains upon exit (Cumming et al., 2004). Further, Mitton (2004) established that the suppliers of equity prefer dividend payouts compared to capital gains when they fear expropriation by insiders. From this viewpoint, the findings of Faccio et al. (2001) and Mitton (2004) provide an interesting insight as they found that investors with stronger rights will use those rights to extract dividends from the firm. The position of banks involved in equity financing exposes another important aspect of harvesting in the form of dividends. If they join in the financing of a successful business, they prefer dividend payouts and are happy to keep their principal in the firm. Usually, passive investors with minority stakes realize their investment by relying on annual dividends simply because they fail to define a specific exit route in advance, as Wall and Smith (1998) showed for the European case.

The last very important factor affecting the choice of these harvesting forms is high growth, both the growth a firm was able to produce in the past as well as future growth potential. Timmons (1999) was among the first to point out that dividend payouts can be regarded as a harvesting choice by the owners of firms that have had fast growth to date but which have matured and reconciled their businesses. Kensinger et al. (2000) noted that paying out dividends prevents the cash flows of a mature business from being reinvested in negative NVP projects. In these circumstances, investors harvest their investment by withdrawing the cash and reinvesting only that amount of cash needed to maintain current markets (Petty et al., 1999). Such firms are characterized by large free cash flows which do not force the current owners to withdraw and attract investors. These firms needed to have the capacity for debt and reinvestment and therefore had to be characterized by a high margin profitability (Timmons, 1999). Sometimes, the choice of harvesting in the form

of free cash flow and dividend payouts also provides investors and entrepreneurs with a means to credibly testify to the firm's good standing to outsiders. Low et al. (2001) thereby claimed that investors in small firms understand the dividend decision as a function of both bank monitoring and signaling in capital markets.

Even Petty et al. (1999) and Timmons (1999) warned that firms with such a harvesting choice must have a long-run earning potential and some opportunities to grow despite the likelihood of them being mature ventures. Tajnikar and Došenovič (2003) found in the case of high-growth firms in Slovenia that investors in mature firms opt for dividend payouts in circumstances where the existing organizational form does not constrain future growth. The long-run earning potential can thus in particular be achieved if they engage in less risky ways of doing business. This usually implies that the firm's markets and its long-term business are associated with lower risks by pursuing a strategy of maintaining current markets and not trying to expand the present markets or expand into new ones (Petty, 1994; Tajnikar and Došenovič, 2003). The longevity of returns to investors in this form can be ensured by investor protection tactics and indeed the ownership structure. Provided that the possible difficulties of these mature firms can be resolved without liquidating the firm and that the yields satisfy investors' expectations, according to Tajnikar and Došenovič (2003), strong owners will opt for dividend payouts. La Porta et al. (2000) thus found that dividend payouts are higher in countries with the stronger legal protection of minority shareholders. Mitton (2004) generalized this finding by asserting that country-level investor protection and firm-level corporate governance are determinants of dividend payouts.

However, from the viewpoint of prospective future growth this group of harvesting forms means "intentionally limiting the firm's growth by paying out operating cash flows to the owners instead of reinvesting them in the business" (Petty et al., 1994). The danger of habitually drawing out cash in the form of dividends is, of course, that valuable growth opportunities that come along are passed over, leading to "an unintended reduction in harvestable value" (Petty et al., 1994). In the case of countries with strong investor protection, La Porta et al. (2000) found there is a strong negative relationship between growth opportunities and dividend payouts. Mitton (2004) confirmed this finding at the firm level. This implies that even investors with stronger rights may not prefer higher dividends if they believe the firm has good investment opportunities.

### *3.2. Public Offerings and Initial Public Offerings*

The term public offering implies that a significant portion of a firm's equity is sold in the public market. In an initial public offering—IPO—a firm issues shares to be sold to members of the public for the first time. As Petty

et al. (1999) explained, IPOs are primarily a means of raising growth capital and only secondarily an exit option of the firm's founders. Prasad (1990, 1994) and Prasad et al. (1996) thereby differentiated between three types of IPOs. In the case of pure primary offerings, only the firm offers the shares to the public; in the case of pure secondary offerings, some of the existing shareholders are exiting the firm and offer some or all of their shares to outside investors in the public offering, while in simultaneous primary and secondary offerings both new shares and the shares of some exiting shareholders are simultaneously offered to outside investors in the same public offering. We can, of course, consider an IPO as a form of investor exit and therefore a form of harvesting only in the second and third cases of Prasad's classification. Cumming and MacIntosh (2003a), moreover, pointed out the important difference between full and partial IPO exits. A full exit IPO involves the sale of all of the investor's holdings for cash within one year of the IPO. Conversely, a partial exit IPO involves the sale of only part of the investor's holdings within one year. Mikkelsen et al. (1997) and Pagano et al. (1998) found evidence that IPOs indeed frequently lead to a harvest. Start-up firms tend to go public to finance an expansion while established companies go public to liquidate the owner's shares (Mikkelsen et al., 1997). Even if the owners choose to separate the event of harvesting from the IPO, deciding to divest at a later time, flotation still offers them greater liquidity and facilitates the eventual harvest of their investment, according to Zingales (1995).

CEOs favor IPOs of all the harvesting forms. An IPO is considered the most desired harvesting method by 65% of CEOs according to Holmburg's (1991) survey, and following an IPO they commonly express satisfaction with this decision (Desroches and Belletante, 1992; Desroches and Jog, 1989). Isaksson (1998) showed, for example, for Swedish venture capital firms that IPOs were the most preferred exit mechanism in the period from 1993 to 1998. Bygrave and Timmons (1992) pointed to one of the primary reasons for this when they established that IPOs generate almost five times greater profits than the second most profitable harvesting form—acquisitions. Notwithstanding the IPO being a desired form of harvesting for both the owners and managers, there are numerous contributing factors for opting for an IPO. We can find them among all four groups of factors affecting the choice between various forms of harvesting. First, analyses of IPOs show that on one hand these factors reflect the position of the owners of the firm. Second, they also relate to the firm's internal characteristics and, third, to those from the environment that influence the firm's operations. Finally, harvesting in the form of an IPO is also a typical harvest that is formed in dependence on the firm's growth—past attained growth and future growth opportunities.

Starting with the last group of determinants that lead to harvesting in the form of public offerings, Timmons (1990) was thereby one of the first to assert

that IPOs are most appropriate for rapidly growing, young, dynamic firms with stable earnings. His argument pointing to the importance of the relationship between a firm's age and fast growth as the two favorable preconditions for executing an IPO was given additional legitimacy by subsequent research such as work by Tajnikar and Došenovič (2003). Fast growth in the past is also an important condition for the choice of an IPO according to Pagano et al. (1998) who established that firms tend to go public following a period of expansion.

Pagano et al. (1998) confirmed that the likelihood of an IPO is positively related to the firm's size, which is also related to a firm's age. Since an IPO is a matter not only of choice but also subject to proper timing, Giot and Schwienbacher (2003) were even more precise in establishing that the probability of undertaking an IPO exhibits an inverse U-shaped pattern over time. Venture-capital-backed firms, they found, thus first exhibit an increased likelihood of harvesting through an IPO but, as time passes, fewer and fewer possibilities of an IPO exist. They found that larger committed investment amounts decrease exit times. Nonetheless, IPOs simply fall outside the scope of the viable choices for firms that are too small, a characteristic often associated with young firms. Even Petty (1994) claimed that IPOs require proper sized firms. IPOs are a good harvesting option for larger firms due to the high fixed costs of going public (Smith and Smith, 2000) as later confirmed by Brau et al. (2003).

Holmburg's (1991) study showed that financing future growth is the primary objective for going public. Nonetheless, the studies of Pagano et al. (1998), Rydqvist and Högholm (1995) and Goergen (1998) suggested that the stock market is not necessarily used as a mechanism to finance growth but rather as a way to reduce leverage and rebalance the firm's financial structure. It also enables the owners to cut back on their involvement in the firm. Huyghebaert and Van Hulle (2002) further find that the size of the primary portion, representing the issue of new equity to finance growth, significantly negatively affects the size of the secondary portion which represents the offering of the existing equity owned by those investors choosing to harvest at the time of the IPO. This implies that when the primary portion of the offering is already large, the owners divest only a few of their own shares. If building up resources for future growth lies at the forefront of the intention to float a firm, then the preference of existing owners to harvest may be compromised.

The choice of a public offering as a harvesting form thus importantly relates to a firm's future growth prospects. Timmons (1999) considered an IPO as an attractive option when there is possibility for the firm to grow rapidly in the future either through an expansion of the business in the existing market or its move into a related market. Namely, IPOs provide access to long-term capital and also enable the investor to meet subsequent capital needs (Timmons,



1999). Tajnikar and Došenovič (2003) explained that these firms are, however, not usually exposed to high market risks as in such cases this problem is more easily resolved through acquisitions, mergers or takeovers. They found that for those firms more likely to harvest in the form of an IPO there is also little prospect of a failure. In this perspective, Reuer and Shen (2003, 2004) noted that IPOs can be regarded as the opening part of a more extended merger and acquisition process. IPOs are often strategic choices rather than simply financial decisions.

Rock (1994) added that a firm can use an IPO as a way of preparing for a subsequent sale. IPOs can thus also reduce the *ex-ante* transaction costs attending acquisitions by raising the visibility of firms and reducing the search costs and information asymmetry related problems associated with mergers and acquisitions. These findings corroborate Timmons's (1999) indication that IPOs enhance the public awareness of a firm. The latter is particularly important when firms consider an IPO as an attractive option to pay off debt obligations (Smith and Smith, 2000). From this viewpoint, it is important to note that compared to other harvesting forms this group provides higher valuations of the investors' holdings (Petty, 1994). Normally, existing investors do not sell their shares during an IPO, according to Smith and Smith (2000). In that way, they signal the firm's good standing and positive future earning prospects. Still, as found by Prasad et al. (1996), for the exiting shareholders of a small firm, the best time to harvest their investment may be exactly at the time the firm goes public for the purpose of raising additional funds for expansion, pay off debt, etc. In these circumstances underpricing is less probable.

Indeed, the problem of asymmetric information determines the choice between a full exit IPO and a partial exit IPO in an important way. The greater the degree of information asymmetry between the seller and the buyer, the greater the likelihood of a partial exit in order to signal quality (Cumming and MacIntosh, 2003a; Gompers and Lerner, 2004; Leland and Pyle, 1977; Lin and Smith, 1997). A partial exit IPO can just as well enable a gratifying harvest for an investor and at the same time reduce the aforementioned information asymmetry and bring about high yields for the investor along with establishing an appropriate financing structure to finance the future growth of a firm. In addition, Gompers and Lerner (2004) pointed to a partial exit IPO as a commitment device to alleviate moral hazard problems and as a mechanism for underwriters to extract additional compensation for the issuing firm.

Apart from IPOs being best suited for young and fast growing firms with ample future growth prospects, firms opting for an IPO should also be able to demonstrate stable earnings (Timmons, 1990) and requisite management (Petty, 1994). In line with Bygrave and Timmons (1992) and Wall and Smith (1998), and in contrast with Cumming (2002), an IPO's internal rates of return are higher than those of trade sales. With a certain degree of skepticism, Bienz

(2004) adds to Timmons's findings that highly profitable companies are more likely to go public, while less profitable ones will be sold. Yet caution is necessary when relying on this finding. When comparing the prices of IPOs and trade sales, one has to take into consideration that firms undertaking an IPO are generally star performers, whereas firms that go through a trade sale also include those unsuitable for going public (Petty et al., 1999). Bienz (2004) believes that findings based on direct comparisons may be misleading in empirical research that does not take selection bias into consideration. In his opinion, highly profitable firms require less control and this allows for an IPO with passive shareholders. Less profitable companies need more control, however, and in such circumstances the original owners sell the whole firm to new owners who take full charge of the firm.

Based on the findings outlined above, another issue particularly important to public flotation compared to other groups of harvesting forms is the influence of the relationship between the firm's management, owners or investors and entrepreneurs. For an IPO, the firm must have requisite management which is willing to dedicate time, effort and financial means to the process of going public, according to Timmons (1999). Yet analyses have shown that the role of the investor or owner proves to be even more important. IPOs are often part of a larger process of transferring control rights in organizations (Mikkelson et al., 1997). From the perspective of the future governance structure of a firm, according to Bienz (2004), an IPO with passive shareholders is a sound choice if less control is required. If, on the contrary, firms need more control, then original owners sell the whole firm to new owners who take control of the firm rather than going for an IPO. From the perspective of a firm's existing governance structure, Brau et al. (2003) empirically found that the probability that the firm conducts an IPO is positively related to the percentage of insider ownership.

Combined, these last two findings show that firms opt for an IPO under conditions where there are numerous insider owners who see the IPO as a vessel for harvesting, and if either the firm's management or owners believe that the firm will not need owners to exert much control in the firm's future. For that very reason, according to Isaksson (1998), firms with an IPO strategy also from this perspective tend to be more active with contacts and analyses of potential buyers than the average venture capital firm.

Studying the behavior of owners, Meulbroek (2000) found that in highly volatile Internet-based firms many owner-managers decided to diversify their portfolios. Huyghebaert and Van Hulle (2002) found that owners of larger firms generating abundant cash flows divested a significantly larger fraction of their own shares at the IPO. This result is surprising as one would expect that owners would feel the need to diversify to a larger extent in the case of firms that are relatively younger and smaller with considerable debt or limited internal cash flow generation. This finding suggests that diversification is not the

main motive for including the secondary portion of stock, this being the stock already held by the present owners in the offering and which represent a harvest by the existing owners. This conclusion is in line with Pagano et al. (1998) and Goergen (1998) who found that IPOs are mainly used to reorganize the ownership structures of firms. In larger firms that generate sizable internal cash flows, the firm-specific investment of initial owners will be less essential which makes the selling off of shares more attractive. Zingales (1995), for example, pointed out that going public can maximize the proceeds in a later sale of the firm. By selling off a minority stake to a widely dispersed shareholder base thereby reorganizing the ownership structure, owners may be able to increase the surplus they can extract from the future buyer.

Moving on to the importance of the entrepreneur, however, Bienz (2004) and Schwienbacher (2005) established that it is impossible to adequately explain an exit choice without taking into consideration the benefits of the venture capitalist (investor) and entrepreneur in the post-exit period. This is the crucial question when deciding on the form of harvest chosen by the entrepreneur or a venture capitalist. When comparing a trade sale and an IPO, they pointed out that in the case of an IPO the entrepreneur retains their equity stake. Where a trade sale takes place, the original owners sell out and the new owner replaces the incumbent management. From the viewpoint of motivation, the entrepreneur's motives are not solely financial. The entrepreneur's private benefits with an IPO can exceed the mere economic benefits from a private sale of shares. These nonmonetary factors can compensate for the entrepreneur's risk of losses incurred through an IPO. The entrepreneur will at times opt for an IPO in spite of the fact that a venture capitalist would choose to sell shares in a packet. This is in line with Cumming's (2002) finding that IPOs are more likely and write-offs are less likely when entrepreneurs control the exit decision.

The connection between the entrepreneurial role and the investor role is very illustrative for the specific position of business angels. It explains why it is less common for business angels to harvest in the form of IPOs. Business angels use IPOs solely for harvesting in high-performance investments (Mason and Harrison, 2002). It is interesting to combine this with the findings of Erikson and Sørheim (2005) who found that technology angels have different harvest expectations compared to other informal investors. They most often harvest through sales to institutional investors and IPOs. Jaaksson's position (1998) on the subject is equally relevant and inciting, being in line with the general findings on the relationship between entrepreneurs and investors. In the case of publicly owned venture capital firms in Sweden in the period between 1993 and 1998, he found that none of the studied firms opted for an IPO.

Environmental factors also play a significant role when deciding for a public flotation. The most important is the condition of the stock markets. Choe et al. (1993) and Bayless and Chaplinski (1996) showed that, in periods of "hot"

markets characterized by less uncertainty and reduced levels of asymmetric information, windows of opportunity for IPOs are more easily created. Black and Gilson (1998) also supported this by saying that active stock markets facilitate IPO exits. They came to support Lerner's discovery (1994) that venture capitalists are more likely to take companies public when their valuations are high, especially in the case of market peaks. MacIntosh (1997), Pagano et al. (1998), Cumming and MacIntosh (2003b, 2004) and Cumming et al. (2004) in their empirical work all found that firms with higher industry market/book ratios are more likely to go public, which is consistent with Gompers and Lerner's (2004) seminal work on U.S. venture capital markets. This can be expected since higher industry market/book ratios indicate higher growth options in the spirit of Fama and French (1992). Cumming's (2002) finding that IPOs are more likely when there are a greater number of syndicated venture capital investors may perhaps also be related to the fact that greater numbers of willing investors can be found in markets when they are flourishing. Provided a firm is opting for an IPO, Giot and Schwienbacher (2003) found that the syndicate size (meaning the number of investors) does not affect the investment duration or actual timing of the IPO. But if the firm's owners and management still have to decide whether the IPO is the best option, this may be a relevant factor. Namely, IPOs call for better firm recognition among potential investors, which may acquire information on new potential investment through investor networks and firm alliances. Indeed, research by Cumming (2002) indicates that IPOs are also more likely when there are a greater number of syndicated venture capital investors, whereas the work of Stuart et al. (1999) showed that firms with alliances also tend to go public sooner. Pagano et al. (1998) concluded that a stock market boom is a necessary but not a sufficient condition for an IPO wave. Huyghebaert and Van Hulle (2002) found that when market conditions start to cool off, past stock market returns still positively relate to the amount the existing owners are able to cash in on from selling their shares in the stock offering. When the cost of IPO underpricing becomes relatively large, however, owners may try to maximize their proceeds by limiting the number of their shares sold at the IPO itself. They would rather sell their shares gradually over time when more information becomes publicly available.

Cumming and MacIntosh (2003a) and Cumming et al. (2004) empirically showed that the broader impact of different legal and institutional factors across countries also has important implications. For example, while IPOs and secondary sales are more likely to be effected as partial exits in Canada, this is not also the case in the U.S. Let us not dismiss the importance of environmental factors for the decision-making process leading to a choice of an IPO, as already indicated by Petty's (1994) finding that firms are in a position to go public if they are in the right industry. Reuer and Shen (2003, 2004) confirmed this view by saying that a sequential divestiture through IPOs is more likely in

industries with spatially-dispersed firms, for high-tech companies, in R&D-intensive industries and for firms with significant intangible resources that have not previously engaged in alliances. Similarly, the empirical results of Brau et al. (2003) confirmed that the probability a firm conducts an IPO is positively related to the concentration of the industry, a high-tech status and the “hotness” of the IPO market relative to the acquisition market.

### *3.3. Trade Sales, Buybacks, Management Buyouts, Employee Buyouts and Management Buy-Ins*

A trade sale occurs when a third party buys the equity shares from existing investors. A buyback is one way for existing investors to exit the investment. In the case of a buyback, the investor's share of equity is repurchased by the entrepreneur, other insiders of the firm and/or the company. A buyback can occur as a management buyout (MBO) in which the existing managers are buyers. Similarly, an employee buyout is an example of a buyback through which employees purchase the equity share held by the investor. This type of exit is often studied in the form of ESOPs (employee stock ownership plans). A trade sale can also occur in the form of a management buy-in (MBI) through which outside managers purchase the equity share held by the current investors. All these transactions involve the exit of all or some investors and thus also represent their harvest. The common characteristic of all these transactions is that they can be either outright cash transactions or a means to form strategic partnerships with the new owners. In addition to the above mentioned group of harvesting forms, King (2002) highlighted one particular case of smaller firms and notes that the Internet boom contributed to a new type of exit through which the owner-founder sells up but stays on as part of a larger organization. King's “selling up, not out” is not of a strategic nature from the perspective of the seller. It is, however, of strategic importance to the buyers.

According to Black and Gilson (1998), Wall and Smith (1998), Fleming (2002), Giot and Schwienbacher (2003), Cumming and MacIntosh (2003a, 2003b) and Schwienbacher (2005), a trade sale exit is the second most preferred form of harvesting following the IPO in both Europe and the U.S. Even though trade sales are not the most preferred harvesting form, Wall and Smith (1998), who examined venture capital exits in Europe, showed that by far the largest number of exits does take place in the form of a trade sale and they reported buybacks as the second most common exit route. IPOs therefore emerge less often than trade sales despite the fact they are the most preferred form of harvesting. Namely, an IPO is carried out in very specific favorable conditions that do not occur often in the firm. For trade sales, on the other hand, the window of opportunity extends over a longer time period and the selection of

a trade sale is also less defined by the type of firm and its internal and external characteristics.

In comparing IPOs, one can thus expect greater heterogeneity in both types of firms doing a trade sale and the conditions in which trade sales occur given that firms undertaking a trade sale as opposed to an IPO are not necessarily star performers. Mason and Harrison (2002), for example, examined the behavior of business angels and established that trade sales were used for harvesting both investments with a good performance and those with lower IRRs and even for investments that only broke even. Bienz (2004), on the other hand, concluded that highly profitable firms are more likely to go public, while the less profitable ones will be sold. Tajnikar and Došenovič (2003) also showed that harvesting in the form of a trade sale is more likely to occur in circumstances of low business performance that does not meet the expectations of existing owners. These findings are further supported by Giot and Schwiendbacher (2003) who established that a trade sale is a type of exit also available to less successful start-ups. They also showed that there is a slight increase in trade sales with the increasing stage of a firm's development. This is in line with the conclusion of Cumming et al. (2004) that buybacks are more likely for firms without significant growth prospects considering that this harvesting form brings no additional capital into the firm. Tajnikar and Došenovič (2003), on the other hand, also showed that investors decide for harvesting in the form of a trade sale when the firm in its current organizational form still has the capacity to grow but has weaker and dissatisfied owners. Cumming et al. (2004) further found, somewhat surprisingly, that larger investments are more closely connected with private exits than IPOs. This clearly shows that the size of the investment is among the important factors determining the choice between an IPO and a trade sale and it thus has to be taken into consideration alongside the firm's performance and growth prospects. Trade sales are also quicker and easier than IPOs especially considering the difficulties in finding a lot of obvious buyers (Wall and Smith, 1998). Wall and Smith (1998) thus pointed out that buybacks or redemptions by co-investors or management are a common exit route for passive inventors and all other investors when other routes fail.

In addition to the firm's internal characteristics discussed above, the effect of the firm's environment also has to be taken into account when examining the probability of trade sales. Cumming et al. (2004), for example, provided evidence that private exits and also write-offs are statistically more likely in countries with smaller stock markets. It is, however, true that also in this case a large enough pool of corporate contacts is needed for trade sales to occur. Giot and Schwiendbacher (2003), for example, showed that a larger syndicate increases the pool of corporate contacts required to find a buyer and consequently do a trade sale. In addition, Amit et al. (1998) found that trade sales are an appropriate way to exit in circumstances of significant

informational asymmetries in which it is too difficult to sell shares in a public market where most investors are relatively uninformed. In the circumstances of greater information asymmetries, many exits take place through sales to informed investors such as other firms in the same industry or to the firm's own management. Cummings and MacIntosh (2003a) also emphasized in this case that one advantage of this group of harvesting forms is that it can involve the sale of only part of the investor's holdings thus enabling only the investor's partial exit. The partial exit signals quality, especially in circumstances of a high degree of information asymmetry between the buyer and the seller.

There are other environmental factors besides market characteristics that determine the probability of trade sales. Cumming et al. (2004), for example, found that a lower proportion of private exits (acquisitions, secondary sales and buybacks) is observed among countries with high legality indices and that buybacks are more common among countries with worse legality indices. In addition, Kaplan (2003) noted that in a weak economy selling a business may be risky for entrepreneurs.

Even though Kaplan (2003) considered a trade sale as a means to increase the value of the venture by selling an equity stake to a strategic partner, there are certain cases in which no strategic goals are ascribed to making a trade sale. That is why several authors like Petty et al. (1999) distinguished between strategic and financial sales.

A financial trade sale as a harvesting form can emerge either when the existing owners realize their inability to continue to control the firm or when they believe their role has been fulfilled. Tajnikar and Došenovič (2003) also showed that a firm seeks new owners when its business performance does not meet the expectations of the existing owners who wish to transfer their ownership to new owners that can bring about the necessary changes. In these cases, existing owners do not consider forming a strategic partnership as a means to resolve the existing problems but look for a new owner to assume the role of a financial buyer. Namely, as already mentioned Bienz (2004) concluded that for highly profitable companies less control is required and this allows for an IPO with passive shareholders. Less profitable companies, on the other hand, need more control and in such circumstances original owners sell the whole firm to new owners who take control of the company. It is also important to note that such buyers are willing to pay more as all and not only part of the equity is available and control of the company is a valuable asset to acquire. It is therefore more difficult to sell a minority interest because it involves selling control (Gladstone and Gladstone, 2004). A trade sale is also more difficult when the top management is unwilling to give up control of the firm they have helped to create or grow (King, 2002). Namely, where a trade sale takes place, the original owners sell out and the new owner usually replaces the incumbent

management. Changes made to the firm's operations may also impose a high cost on firm's employees (higher pressures, layoffs, etc.).

In the case of financial sales, the new owners feel that the firm can be operated more effectively in the future or that the firm is worth more "dead than alive." In the latter case, a financial sale involves either withdrawing the firm's free cash flows over time and maintaining the status quo thereby liquidating the firm's assets gradually or liquidating the firm's assets immediately. In these circumstances, financial buyers look for a firm's stand-alone cash-generating potential as the source of value. The above conclusions are also in line with the findings of Mason and Harrison (2002) showing that exits from "living dead" investments have been primarily through sales to other shareholders or third party investors.

There are also several reasons for financial sales from the viewpoint of sellers. Petty et al. (1999), for example, showed that such a sale enables investors to diversify their investments. In addition, Petty et al. (1999) showed that the most common reason for selling the firm relates to estate planning. The analysis of family firms also indicates that trade sales are not used to pursue strategic goals. Upton and Petty (1998) thus showed that providing liquidity for the exiting family members, obtaining growth capital and maintaining family control are the three objectives pursued by trade sales in the case of family firms. Private capital can also be infused to help a family-controlled firm transfer ownership from one generation to the next while providing growth capital. Where family businesses are involved, private equity placement is used as an alternative to an IPO. During the 1980s, financial sales also occurred in the form of so-called bust-up leveraged buyouts (LBOs) (Petty et al., 1999). In cases of bust-up LBOs, the new owners sold the acquired firm's assets off. Bust-up LBOs were replaced by build-up LBOs in the 1990s. These involve constructing a larger enterprise that is then taken public via an IPO. The newly formed combination is operated privately for five to seven years in order to establish a track record of success and is then taken public. In the latter case, trade sale was a means to create conditions suitable for the optimal form of harvesting, that is, an IPO.

That trade sales are not used to pursue strategic goals holds true for both sellers and buyers. Trade sales are thus likely to be unattractive when corporate venture capital firms invest for strategic reasons. In some cases, however, a trade sale brings new owners into the firm who are interested in more than the firm's stand-alone cash-generating potential and envision both efficiency improvements and synergies that can enhance the value of the combined firms. In this case, trade sales can have the nature of strategic sales. If a trade sale does occur for strategic reasons, strategic buyers often pay a higher price than purely financial buyers (Petty et al., 1999). That is why buyers in the case of such trade sales differ from buyers in the cases of IPOs. MacIntosh (1997) established that



strategic buyers are larger established companies (industrial buyers) or another venture capitalist. Erikson and Sørheim (2005) also mentioned institutional investors as such buyers. Further, according to King (2002), the buyer is often a rival business or customer that knows how to manage in the sector.

Strategic sales and purchases for this group of harvesting forms occur for several reasons. A strategic purchase is often made by public and private companies “that see their prime means of growth as purchasing other businesses” (Gladstone and Gladstone, 2004). It is also often made by larger corporate buyers interested in buying a small business that has made significant progress and has new ideas. This is beneficial for the small firm as the corporate partner knows more about the marketplace and the production of the product. It is also attractive to the corporate partner as it gives it the possibility to own the entire firm in the future (Gladstone and Gladstone, 2004).

*3.3.1. Management Buyouts and Management Buy-Ins* Management buyouts (MBOs) “involve incumbent managers in acquiring a significant equity stake.” Management buy-ins (MBIs), on the other hand, “involve external managers, as individuals with institutional support, acquiring control of the company” (Wright et al., 1993). MBOs are possible when the existing partners or key managers are interested in buying the business (Kaplan, 2003; Timmons, 1999). They are interested in buying the business when the price of the firm is low and when the possibility exists for managers to reap successful past growth more as owners (Stewart, 1993). The lower price of the firm in this case is the consequence of the firm’s fast past growth “showing very little profit along the way” (Timmons, 1999). Managers can achieve a more successful harvest than existing owners if, according to Stewart (1993), incentives are needed for the managers to act more like owners. It is precisely MBOs that are a good way to create such incentives (Petty, 1994). In such cases, managers are stimulated to participate in the purchase due to optimistic expectations regarding future growth that can be achieved through their increased motivation and involvement. The strength of the management team and the business’ ability to support debt financing (the level and stability of operating cash flows, the value or tangibility of assets as collateral, the extent and nature of the firm’s growth opportunities) are thus clearly key factors that determine the selection of an MBO as a form of harvesting (Petty et al., 1994).

The desire to pursue MBOs can also be part of the firm’s strategy. King (2002), for example, showed that firms divesting themselves of non-core subsidiaries to focus on their core business support this form of harvesting as MBOs and MBIs allow entrepreneurial management teams to buy their non-core divisions. In these circumstances, an MBO is often a means to “unleash new entrepreneurial skills,” improve the performance of the business and create smaller, more flexible firms. Such buyouts can occur only if there is

sufficient financial support provided by leveraged buyout firms that specialize in supporting entrepreneurs and financing the purchase of established companies and underperforming divisions of large firms.

Giot and Schwienbacher (2003) highlighted another important aspect of harvesting in the form of an MBO and an MBI. They state that the selection of MBOs and MBIs as harvesting forms is also explained by the logic of information asymmetry and adverse selection. These asymmetries are prevalent in the most volatile industries where the variability of firms' returns within the industry is great. That is why MBOs occur more often in such industries. In the case of MBOs, fewer information asymmetries exist between buyers and sellers (Howorth et al., 2004). That is why MBOs progress more easily than other forms of harvesting in cases when the firm's current situation is appropriate for an MBO. Howorth et al. (2004) highlighted another specific feature of MBOs. Within the agency theory framework, the authors found there are fewer information asymmetries and knowledge transfer is more easily facilitated where the MBO (or MBI) is part of the family firm's long-term strategy. The desire to resolve information asymmetries also probably explains why a combination of management buy-ins and buyouts (BIMBO) is used as a harvesting form. Where an MBO is combined with an MBI the team buying the business includes both existing and new managers (King, 2002).

*3.3.2. Employee Buyouts* Most authors discuss ESOPs (employee stock ownership plans) when they explore employee buyouts as a form of harvesting. At least half of all ESOPs are used to create a market for the shares of exiting owners of closely held firms (Petty et al., 1994). This clearly shows that ESOPs can also be used as a form of harvesting. For Petty et al. (1994), three key factors determine the selection of an ESOP as a form of harvesting. First, a firm must be able to support high levels of debt. Second, the entrepreneur must be willing to relinquish control to the employees. Third, for employees this type of harvesting means that both their jobs and their retirement funds depend on the success of the business. That is why it may not be suitable for all firms. It could, however, be successful for industries with chronic overcapacity and intense management-labor conflicts. Ding and Sun (2001) also believed that three main reasons are typically offered to explain ESOPs. These include tax advantages, a reduction in financial reporting costs and a reduction of agency costs. Kensinger et al. (2000) add that in some cases a sale to employees results from the entrepreneur's wish to see the firm continue with as little change as possible after the founder's role is reduced.

Despite the fact that several reasons are typically offered to explain ESOPs, there is a general consensus in the literature that employee buyouts are characteristic of companies where there are conflicts between the owners and employees and where they can be used as a tool to resolve these conflicts

(Petty et al., 1999). Namely, agency problems are reduced when there is less divergence of interests between agents and principals (Jensen and Meckling, 1976). In such circumstances, employee buyouts serve as a positive motivational device for employees (Timmons, 1999). Evidence from the Singapore economy demonstrates that ESOPs align managerial with shareholders' interests and contribute to company performance (Ding and Sun, 2001). In addition, Matsunaga (1995) reported that the lower the value of reported relative to target income the greater the value of ESOPs per employee issued. Contrary to the above mentioned authors, Yermack (1995) found only weak support for the relationship between agency-cost reductions and ESOPs for U.S. firms.

It is, however, important to note that the selection of this harvesting form is not only determined by the presence of conflicts between managers and employees. Certain other firm characteristics are also important. This form of harvesting is mostly attributed to closely held companies (Timmons, 1999). Petty et al. (1999) highlighted that employees might be the ones most interested in buying the firm through an ESOP if the firm is not a good candidate for an IPO or if no strategic or financial acquirer can be found. The motive of buyers in these cases is to preserve employment and enjoy the benefits of ownership. The empirical analysis of Ding and Sun (2001) further showed that in Singapore ESOPs are adopted by firms with a large size, high growth potential, a low times-interest-earned ratio (a proxy for the debt-servicing capacity) and lower liquidity. However, their actual data showed that, among the firms that actually use ESOPs, larger firms tend to use proportionally less.

### *3.4. Acquisitions, Mergers and Takeovers*

A merger is a deal characterized by a one-for-one share swap between the merging firms; if the swap is not on equal terms then this is an acquisition. An acquisition is therefore any deal where the bidder ends up with 50% or more of the target (King, 2002). Any action where an acquiring firm makes a bid for the acquiree is usually referred to as a takeover. Kaplan (2003) considered a merger as an alternative to selling an equity stake to a strategic partner and similarly, Cumming and MacIntosh (2003a) stated that in the case of an acquisition the buyer is typically a strategic acquirer. This group of harvesting forms is thus typically associated with creating strategic partnerships.

Even though forming strategic partnerships is a common characteristic of acquisitions, mergers and takeovers, they can be differentiated according to their underlying cause. In the literature we can find at least six groups of reasons for forming strategic partnerships. The need to form strategic partnerships through acquisitions, mergers and takeovers emerges when the firm experiences problems, when expectations are optimistic, when additional capital is needed and the current owners cannot provide it, when they cannot be prevented

by existing owners, in cases of expected synergies and in circumstances of external shocks. It is important to note that these reasons for creating strategic partnerships are not related to correcting managerial failure. This conclusion is supported at least to some extent by Franks and Mayer (1996) who did not find strong evidence that hostile takeovers are motivated by the correction of managerial failure.

The most profound reason for forming strategic partnerships through acquisitions, mergers and takeovers is certainly the firm's loss of competitive advantage and consequently the highly probable occurrence of significant difficulties in running the business (Timmons, 1999). As also shown by Kaplan (2003), a merger should be considered when a company loses its competitive advantage in the marketplace, when it needs to protect its business position (e.g., by protecting itself from an unwanted takeover, gaining a patent position in the market, etc.), when it is barely surviving (e.g., due to a low cash position, technology that is becoming obsolete, loss of market share to competitors, management that is leaving the firm, etc.) and when the firm needs to acquire added services (e.g., by acquiring new management talent, entering new markets and gaining financial resources). Tajnikar and Došenovič (2003) further showed that investors decide to exit through acquisitions, mergers or takeovers when faced by high-risk markets. The acquired firms are usually less profitable (Cosh et al., 1984; Levine and Aaronovitch, 1981; Meeks, 1977; Singh, 1971, 1975). Thompson (1997) also found profitability was negatively related to the probability of a takeover for U.K. Building Societies between 1981 and 1993. In Germany, private and public corporations are also more likely to be acquired or to fail when their performance is poor (Altman, 1968; Denis and Sarin, 1999; Maksimovic and Phillips, 2001) and leverage is high (Powell, 1997; Zingales, 1998). In addition, Shleifer and Vishny (2003) found that acquired firms have equity that is undervalued or relatively less overvalued. The case of high-growth firms in post-transitional countries, however, does indicate that business performance is not necessarily a key determinant of this group of harvesting forms. Tajnikar and Došenovič (2003) showed that in Slovenia such firms achieved results that met the expectations of their strong owners and their organizational form did not pose a constraint on further growth. Their conclusions imply that current owners fear that future conditions in risky markets may be unfavorable to the firm's future performance. Interestingly, Dickerson et al. (2002) found that it is current profitability—and not pessimistic expectations for the not-so-near future—that is crucial to takeover actions. Their study of the experiences of British firms in the 1970s and 1980s showed that firms without positive NPV investment opportunities did not experience a significantly increased takeover hazard if they increased investment or reduced dividends contrary to the prediction of the free-cash-flow theory (Jensen, 1986).

Cumming (2002) contributed to this finding by showing there is no evidence that market/book differences affect the likelihood of an acquisition.

Acquisitions, mergers and takeovers, however, can also be stimulated by optimistic expectations of future developments. Andrade and Stafford (2004), for example, found that mergers are positively related to sales growth and profitability. Through mergers, companies increase their capital base in response to good growth prospects when capacities are near to full utilization. Andrade and Stafford (2004) further showed that mergers can also be positively related to capacity utilization if there are growth opportunities at the industry level.

Another reason for the increased likelihood of acquisitions, mergers and takeovers due to expected good conditions is mentioned by Timmons (1999). According to him, this type of harvesting also enables the founders of smaller companies to obtain a substantial amount of capital from a larger company. This is especially true where there is a need for financing growth and the current owner does not have the capacity to provide the funds needed (Petty, 1994). Petty et al. (1994) also noted that merging or being acquired is most appropriate when the selling firm has significant growth opportunities requiring outside equity capital. The studies mentioned above clearly show that both existing and expected circumstances prevailing within the firm and its environment affect acquisition, merger and takeover activity. In the first case, they are encouraged by the existing problems and poor conditions while, in the second, they are stimulated by positive expectations and good prospects.

Whether existing circumstances prevailing within the firm affect the probability of acquisition, merger and takeover activity also depends on the strength of existing owners. Köke (2002), for example, found that, for private and financial owners, poor company performance is not as strong an incentive to sell as it is for nonfinancial owners. He thus found that public corporations under strong ultimate ownership are less likely to be acquired. Strong ultimate ownership is determined by the concentration of ownership and the type of the ultimate owner.

The prevailing reason for acquisitions, mergers and takeovers is the synergies they can create. Mueller (1969) argued that conglomerate acquisitions can be explained by the existence of management synergies, that is, financial synergies and risk reduced due to the pooling of activities. Petty et al. (1999) also showed that forming strategic partnerships is reasonable when the possibility exists for certain synergies to emerge between firms. The conclusions of Cumming and MacIntosh (2003a) that strategic acquisitions often involve the merger of two firms with some prior relationship are thus not surprising. They further showed that the acquiring firm is most commonly a larger entity in the same or similar business as the acquired firm. Kensinger et al. (2000) also showed that, in the case of strategic acquisitions competitors, customers or

suppliers are good potential acquirers with potential synergies from integrating forward and backward. Synergies can also occur with the formation of market monopolies. That is also why market power strengthens the firm's incentive to merge and speed up merger activity (Lambrecht, 2004). It is, however, also true that the firm's size provides a hindrance to a takeover. This is primarily due to the existence of financial constraints (Cosh et al., 1984; Dickerson et al., 2002; Hasbrouck, 1985; Machlin et al., 1993; Rege, 1984). In Germany, for example, private and public corporations are more likely to be acquired or to fail when the firm size is small (Bethel et al., 1998; Harhoff et al., 1998; Mulherin and Boone, 2000). The latter is also consistent with the theoretical prediction of Shleifer and Vishny (1992).

Analysts of acquisitions, mergers and takeovers devote special attention to shocks that occur in the economy or various sectors and also affect the emergence of this type of harvesting. Lambrecht (2004) offered the theoretical explanation of the relationship between merger waves and cyclical product markets whereby mergers are pro-cyclical. Maksimovic and Philips (2001), for example, provided some empirical evidence on the pro-cyclicality of merger waves. Mitchell and Mulherin (1996) contributed to these findings by observing an increase in merger activity in periods of industry or economic contractions for which information asymmetry tends to be higher. Andrade and Stafford (2004), for example, determined that mergers facilitate industry contraction. Through mergers, industries with excess capacities are consolidated. Their findings complement the conclusions of Jensen (1993) who argued that mergers are the principal way of removing excess capacities. Harford (2004), based on the behavioral hypothesis (McGowan, 1971; Rhodes-Kropf and Viswanathan, 2004; Shleifer and Vishny, 2003), explains merger waves as a function of the managerial timing of market overvaluations of their firms. Aggregate merger waves thus occur when market valuations are high relative to true valuations. The high dispersion of stock market valuations also explains merger waves. However, the strong association between merger activity and high stock market valuations does not explain the clustering of takeover activity across industries (Powell and Yawson, 2005). Therefore, specific industry factors need to be considered when analyzing the clustering of merger and takeover activity as also shown by the neoclassical explanation of merger waves. The neoclassical explanation of merger waves dates back at least to Gort (1969) and more recently to Mitchell and Mulherin (1996) and Jovanovic and Rousseau (2002) who explained merger waves through shocks to an industry's economic, technological or regulatory environments. Powell and Yawson (2005) showed that specific industry shocks that cause takeover activity are low growth, the threat of foreign competition and high industry-adjusted stock market performance. Their findings regarding the threat of foreign competition can be at least partly linked to Bjorvatn's (2004) conceptual paper linking economic integration that

strengthens competition and cross-border mergers and acquisitions. Eisfeldt and Rampini (2003) and Harford (2004) argued that these shocks are not enough on their own and that there must be sufficient macro-level capital liquidity to propagate a merger wave. This liquidity causes merger waves to cluster even if industry shocks do not. Cumming et al. (2003) also established that liquidity conditions determine harvesting opportunities in the form of mergers and acquisitions.

### *3.5. Liquidations, Bankruptcies, Sales of Assets and Write-Offs*

In this section, we deal with a group of harvesting forms where the owners' returns do not take the form of dividends or capital gains. Instead, returns are achieved by the sale of part or all of firm's assets. Here we find harvesting in the form of the liquidation of a firm, a bankruptcy, the sale of a firm's assets and write-offs. A write-off involves a situation in which an investor walks away from the investment and writes down the overstated value of the assets on the firm's balance sheets. Reorganization is a synonym for bankruptcy (Gladstone and Gladstone, 2004). A special case can be considered with "living dead" investments, namely nonviable ventures where the investor nevertheless continues to hold shares and keeps employing the entrepreneurial team. All of these cases are normally considered occurrences of a business failure that eventually, if not immediately, lead to liquidation of the firm. Liquidation does not necessarily mean that the business will be shut down immediately. Assets can be sold and replaced with leased assets thereby allowing the business to continue operating after the owner has partially harvested his/her investment. This is called an orderly liquidation according to Kensinger et al. (2000).

The liquidation of a firm and its assets occurs when the opinion prevails that the firm is worth more "dead than alive" (Petty et al., 1999). Pioneer researchers of bankruptcy predictions are Beaver (1966) and Altman (1968). Relevant results have also been obtained by Zmijewsky (1984), Frydman et al. (1985) and Gentry et al. (1987). A firm may find itself in these circumstances either due to the intention of its owners, but more often not. As Wickham (1998) concluded, in these circumstances the expectations of various stakeholders in the business venture fail to be met. Regardless of the fact that the stakeholders of any firm can be very diverse, it is logical to assume that it is the owners who are most influential and instrumental in the realization of this group of harvesting forms. If in certain cases, however, entrepreneurs control the exit decision, write-offs will be less likely, according to Cumming (2002). This is consistent with Petty et al.'s (1999) case studies on entrepreneurs' attachment to their firms and their reluctance to write their firms off. Donohoe (2004) provided us with a circumstantial confirmation of this hypothesis. He studied U.S. firms and found that firms with high levels of inside equity ownership and secured indebtedness

decided on bankruptcy if they had poorer financial conditions compared to firms with similar ownership and indebtedness characteristics. Firms with a high level of outside equity ownership and short-term indebtedness sought a reorganization while they are still in comparatively better financial conditions. Research on the choice of harvesting forms by Tajnikar and Došenovič (2003) reveals that, in the case of Slovenian high-growth firms, liquidation is opted for by weak owners and in situations where difficulties cannot be resolved by new owners or strategic partners. If new owners could resolve the existing problems, they find that harvesting in the form of a trade sale becomes more likely than a liquidation. The ownership structure proves to be a very important factor of choice for this least desirable harvesting form.

Most of the research to date accumulates knowledge on other determinants that are crucial for predicting the liquidation of a firm or its assets, stemming from either the characteristics of firms themselves, or their environments. The question that has mainly arisen is this: Is it possible to use financial data to assess the probability of a firm's bankruptcy? In a study of U.S. local telecommunication firms, Foreman (2003) explained which firms would fail within two years by certain financial ratios such as earnings per share, return on assets, retained earnings to assets, total debt proportion and working capital to sales. He also showed that firms with a higher market-to-book ratio are less likely to fail. His results were later more generally supported by the findings of Pompe and Bilderbeek (2004). They found that virtually every ratio category (e.g., profitability, activity, liquidity and solvency ratios) had some predictive power for bankruptcy. Certain ratios perform similarly with different populations. The cash-flow-to-total-debt ratio, for example, achieved the best overall accuracy for both old and young firms. Yet long before them, Beaver (1966) determined this to be true for large firms. We must apply some caution in using financial data for the prediction of a bankruptcy as Köke (2002) found that acquisitions and failures both tend to be influenced by common factors.

Köke (2002) for the case of German private and public corporations finds that firms are more likely to fail when financial performance is generally poor. Firms with the lowest market/book ratios obviously are written off (Cumming, 2002). Köke (2002) found, in the case of German private and public corporations, that these firms are more likely to fail when the performance leverage is high. Donohue (2004) further studied U.S. firms and found that firms with a high level of outside equity ownership and short-term indebtedness sought reorganization while they were still in a comparatively better financial condition. On that note, Gladstone and Gladstone (2004) stated that reorganization has been used by many small businesses to remove creditors as well as investors. Opler and Titman (1994) found even earlier that a firm's capital structure also has to be considered as a factor of a bankruptcy announcement.



Firm size also proves to be an important determinant of these harvesting forms. Köke (2002) found that firms in his sample were more likely to fail if they were small, which is consistent with the earlier theoretical predictions of Shleifer and Vishny (1992) that larger firms are less likely to fail. Pompe and Bilderbeek (2004) confirmed Köke's finding: the likelihood of bankruptcy is smaller for larger firms. The background providing more insight into this conclusion can be found in the results of Brüderl and Schüssler (1990) who, for example, explained that the decline leading toward bankruptcy is more gradual for larger organizations as they have more resources open to them in bad times.

Moving from the size to the age characteristics of firms, Pompe and Bilderbeek (2004) also found evidence of, for example, Altman's conjecture (1993) that the older the firm the smaller the likelihood of bankruptcy. We may look for an explanation of this in Levinthal (1991) who said that the prior success of older organizations buffers them against failure for a certain time. Pompe and Bilderbeek (2004) however also made the point that the bankruptcy of young firms is more difficult to predict than the bankruptcy of established firms.

Giot and Schwienbacher (2003) interestingly hypothesized that the likelihood of failure and harvest in the form of a liquidation decreases with the development of a business project and with the timing of funds supplied. Certainly, more money means better resources which allows the entrepreneur to develop the project more quickly. Yet, large initial cash inflows may be adversely related to the success of a project due to a lack of prudence. If a constant amount of funds is disbursed over a prolonged time period rather than given upfront, it prevents the initial pursuit of unsuccessful projects by the management. Authors are unable to extend this finding to the relationship between a project's maturity and the probability of liquidation so it is impossible to say that later-stage projects are less likely to be successful and will be liquidated more often. With respect to duration, the research to date only supports the least bold of the hypotheses that write-offs occur after the shortest investment duration (Cumming, 2002) showing that investors have the least patience with the worst performing businesses.

Firm growth is also a good predictor of business failure. Köke (2002) found that, for public corporations, failure is typically preceded by a significantly lower growth rate of employment. Yet it may obviously not always be the case of lagging historical growth rates. In certain conditions, growth that is too fast also relates to business failure. Foreman (2003) found in the case of the U.S. local telecommunications industry that if large working capital is a precursor to overexpansion then greater capital needs also expose the firm to the dangers of failure.

With respect to business performance, in a study of Italian manufacturing firms Becchetti and Sierra (2003) found that the degree of relative

firm inefficiency, estimated using a stochastic frontier analysis, has significant explanatory power in the prediction of bankruptcy. They added to the list of bankruptcy predictors a unique insight that had previously escaped empirical verification—the impact of technical and allocative inefficiency. It seems that, beside the mere financial ratios, it is also necessary to include the reasons for their values showing underperformance.

Becchetti and Sierra's (2003) work supported the findings of Bandopadhyaya and Jaggia (2001) by showing that qualitative regressors such as customers' concentration, subcontracting status, export status and the presence of large competitors in the same region also have significant predictive powers. Bandopadhyaya and Jaggia (2001) examined the case of U.S. firms re-entering bankruptcy. They noted that other information, apart from pure financial data, can also be used to predict a firm's bankruptcy. They found that the probability of a firm re-entering bankruptcy is lower for those firms that take a long time to reorganize, reduce their debt-to-assets ratio, that do not divest and if they belong to an industry that has low capacity utilization and low demand growth.

Finally, two strands of research suggest that characteristics of the environment also significantly affect the occurrence of harvesting in the form of the liquidation of firms and their assets. Giot and Schwienbacher (2003) found there is a greater proportion of write-offs in countries with higher legality indices signifying the greater advancement and efficiency of the legal system. This might be a little surprising, but perhaps it reflects the increased willingness among investors to pursue risky opportunities in those countries with better legal protection. Foreman (2003) suggested that the likelihood of bankruptcy increases for firms relying heavily on "uncertain legal and regulatory outcomes." An older finding that is in line with Foreman is that of Denis and Denis (1995), namely, that the key reasons for financial distress are unexpected macroeconomic and regulatory developments. With respect to equity market conditions, Giot and Schwienbacher (2003) showed that firms were more likely to be liquidated during the stock market bubble years. Internet firms were especially quick to be shut down while, in contrast, biotech firms were the slowest to liquidate, developing Lerner's (1994) assumption that biotech firms mature slowly and do not incur large up-front costs while Internet firms are known to be cash-greedy. The size of the initial cash requirements therefore also seems to be positively related to the probability of liquidation.

#### 4. CONCLUSION

The introduction to this chapter explained that, although the selection of a particular harvesting form is primarily a subjective decision of the investors and owners, the characteristics of the firm and its environment also affect the

selection of a given harvesting form. It was also proposed that certain factors affect investors' and owners' interests and decisions and that the firm's characteristics and its environment are those factors determining both the availability and actual choice of a specific harvesting form. The full range of determinants of harvesting forms can be classified in four groups. These include: (1) the determinants that shape both the interests and decisions of owners and investors; (2) the determinants that influence the conditions within the firm; (3) the determinants that emerge along with the firm's growth; and (4) the determinants that reflect the firm's environment. The authors hypothesized that the factors in these four groups can determine the selection of any of the five groups of harvesting forms discussed in the previous section. Yet the manner and extent to which any group of determinants affects the selection of a particular group of harvesting forms can vary for different groups of harvesting forms. The content and size of each of the four groups of determinants may indicate the relative importance of individual groups of determinants in selecting a specific harvesting form. It might, however, also reveal those determinants that have received insufficient attention by researchers to date. The survey of literature presented in this chapter on harvesting forms and their determinants supports the idea that these determinants can be classified in the aforementioned four groups.

According to the research outlined in the literature survey, harvesting in the form of dividend payouts is often associated with strong ownership, high firm profitability that enables high dividend payouts and stable although not high future growth that does not require the firm to make risky business decisions and thereby increase its probability of failure. It could thus be inferred that the probability of selecting this group of harvesting forms is predominantly affected by those factors pertaining to the conditions found within the firm. Among these determinants, various authors include the firm's market conditions enabling less risky ways of doing business, a mature stage of the firm's organizational development, long-run earning potential and the ability to signal its success to capital markets also through its dividend payouts. Further, the literature survey indicates that the selection of this harvesting form is more probable in the case of strong owners who intend to retain control over the firm, often including entrepreneurs-investors. The existing research also reveals that this harvesting form can be opted for either when firms have a large value, when market valuations are lower, or when firms exhibit high margin profitability. Several authors also discuss determinants associated with the firm's growth. In this respect, a firm opting for this harvesting form can be expected to have had fast growth so far. Further, certain findings indicate a negative relationship between growth opportunities and dividend payouts. Determinants that reflect the firm's environment seem less important. For this group of determinants, one could ascertain that the stronger legal protection of minority shareholders at

the country level has a positive impact on the choice of this harvesting form. To the best knowledge of the authors, no determinants influencing personal interests and the decisions of investors and owners are discussed by the relevant literature. The literature survey thus gives an impression that harvesting in the form of dividend payouts is a fundamental form of harvesting. Investors maintain their investment and gradually harvest it as long as no incentives or constraints emerge to motivate them to withdraw from the firm through alternative harvesting forms.

As opposed to other harvesting forms, IPOs appear to be chosen in very specific conditions. The determinants of IPOs belong mostly to the groups of determinants that shape the conditions within the firm, those associated with the firm's growth and those reflecting the firm's environment. Substantial research points specifically to the importance of factors originating in the firm's environment such as legal and institutional factors and the "hotness" of the IPO market relative to the acquisition market. The latter factor seems to provide one of the answers to the question of why investors decide against an acquisition, a merger or a takeover and in favor of an IPO. Regarding the determinants associated with the firm's growth, IPOs are often shown to be most suitable for rapidly growing firms or firms with fast growth in the past. The latter case explains why IPOs often follow a period of expansion. IPOs have also been found to be motivated by the prospect of fast future growth, or sometimes even as part of an extended merger and acquisition process. The literature review suggests that IPOs are more common in spatially-dispersed, high-tech, R&D intensive industries and industries with significant intangible resources and a higher concentration. Being in the right industry is therefore one of the determinants from the group of determinants that shape the conditions within the firm. Other such determinants include being of a proper size, being young and dynamic, having the requisite management, showing stable earnings stemming from successful growth to date, or needing long-term capital due to the expectation of successful future growth. Also, in the case of harvesting in the form of IPOs, determinants influencing personal interests and the decisions of investors and owners are not particularly highlighted by the existing research. The existing research, however, clearly shows that IPOs are the ideal form of harvesting, yet one that is very restrictive in terms of firm quality and business performance.

However, when the conditions are not best suited for an IPO according to the general consensus, the second best alternative is a trade sale. Poor business conditions or unfulfilled investors' expectations prevent investors from harvesting in the form of both dividend payouts and IPOs. In these circumstances, investors prefer to opt for a trade sale. The literature review thus reveals that harvesting in the form of a trade sale is more likely to occur when the firm seeks new owners in circumstances of low business performance

that does not meet the expectations of the existing owners. This is why in such cases the owners plan to exit their investment and transfer their ownership to new owners. From this standpoint, the key role is ascribed to determinants that shape the conditions within the firm and, among these, to those that relate to the firm's ownership and performance. It has been established that trade sales are motivated by the desire to introduce a new owner to the firm, especially when significant informational asymmetries exist. For trade sales, the performance criteria seems not to be as restrictive as in the case of IPOs, considering that evidence suggest that both well-performing firms and less successful start-ups, less profitable firms or even firms just breaking even adopt this harvesting form. It is, however, important to note that research into harvesting in the form of a trade sale does not indicate a connection to the problems the firm intends to resolve through strategic partnerships, such as market problems, organizational problems and problems related to the firm's financing. That is why trade sales are usually not used to pursue strategic goals. It may be precisely due to the inability of the existing owners to solve the emerging problems within the firm that strong existing owners look for new strong owners to take their place. The strength and great influence of the existing owners probably also explains the conclusion by some authors that such firms are rarely liquidated despite their sometimes poor business performance. Namely, liquidations can to some extent be prevented or postponed by strong owners. That is why existing investors defending their interests encourage this form of harvesting. Contrary to the harvesting forms discussed above, research into trade sales does provide some discussion of the determinants influencing the personal interests and decisions of investors and owners. To be specific, trade sales enable sellers to diversify their investments, they also allow a partial exit over a full one, and facilitate more discretion over the timing of their harvest given that the window of opportunity for trade sales extends over a longer time period. In addition, from the viewpoint of the seller, trade sales may not be used to pursue strategic goals. The opposite, however, can be true of the buyers who may purchase a new firm exactly for the purpose of implementing their previously conceived strategy. Some authors, for example, talk about either the buyer's desire to construct a larger enterprise, or benefits from the envisaged synergies. It is the belief of the authors of this chapter that the determinants reflecting the firm's environment receive no attention in the existing literature.

The latter also seems to hold true for harvesting in the form of a management buyout. For this harvesting form, the existing research identifies determinants that belong to the group of determinants shaping the conditions within the firm. These include having fast past growth showing very little profit along the way, and having a low current price of the firm. One contribution, for example, showed that such circumstances can emerge in volatile industries. A management buyout thus occurs as the most suitable exit route for encourag-

ing managers to act more like owners, thereby increasing their motivation and involvement. In addition, the research notes that higher information asymmetry and adverse selection can prevent the transformation of ownership through other means, thereby prohibiting the selection of alternative harvesting forms, particularly the selection of more general types of trade sales. Evidence suggests that, from the perspective of the relevant determinants, a management buyout can thus be regarded as a type of trade sale.

Yet there is empirical evidence on determinants influencing the personal interests and decisions of investors and owners in the case of harvesting through an employee buyout, as this harvesting form provides both the advantage of benefits of ownership and the benefits of preserving workers' employment. Quite like in the case of IPOs this harvesting form comes into play in very specific conditions. Three key factors determine this harvesting form. The first two belong to the group of determinants that shape the conditions within the firm. The first determinant relates to firms experiencing problems, revealed in their low debt-servicing capacity and lower liquidity, yet possessing certain growth potential. The second determinant refers to being a closely held firm with conflicts between the owners and employees, whereby an employee buyout is a means to reduce agency costs. The third determinant belongs to the group of determinants reflecting the firm's environment. Employee buyouts often seem to be encouraged by tax advantages and possible reductions of financial reporting costs. This harvesting form is another specific type of a trade sale, selected in very specific conditions. Determinants of both employee buyouts and other types of trade sales relate to the entry of a new strong owner and critical conditions within the firm. However, the specific feature of employee buyouts is that the new owners in fact become the firm's employees due to the need to resolve conflicts between the owners and employees.

The literature survey well elaborates the determinants of harvesting in the form of an acquisition, merger or takeover. These harvesting forms are opted for by firms with existing or foreseen problems. Various authors have therefore discussed a large number of determinants that can be classified either as determinants that shape the conditions within the firm or determinants associated with the firm's growth. They find that problems for such firms can emerge in very different ways and for various reasons. According to the research to date, firms selecting this group of harvesting forms experience a deteriorating competitive advantage in the market, are under the threat of foreign competition or face increasing market power concentration. Further, following fast growth in the past their markets may be becoming risky, their capacities could be nearing full utilization and they can be characterized by their low growth, high leverage, poor performance and low current profitability. Some contributions suggest this form of harvesting occurs more often for small firms. They imply that such firms do not pursue harvesting in the form of trade sales since they lack

strong owners. They do, however, seek to benefit from management synergies and other benefits of strategic alliances through an acquisition, merger or takeover. Strategic partners are often selected conditional on prior relationships between the merging firms. The weakness of existing owners is revealed in their incapacity to provide the funds needed. The research shows this form of harvesting sometimes occurs in smaller firms aiming to obtain a substantial amount of capital from a larger firm. This type of harvesting can also be selected as a response to certain adverse conditions in the firm's environment such as shocks to an industry's economic, technological or regulatory environment in periods of industry or economic contractions. What is also characteristic of these harvesting forms is that they can be selected by either individual firms within various sectors or may even cluster within and across industries. Regardless of the wide range of findings, one can ascertain that the basic reason motivating the selection of these harvesting forms are problems the firm intends to solve through strategic partnerships. In the absence of such problems, a trade sale would be more appropriate.

Liquidation is a harvesting form opted for by weak owners and for firms with difficulties that cannot be resolved by new owners or strategic partners. If new owners could resolve the existing problems then harvesting in the form of a trade sale would be more likely. If there were a need to form strategic partnerships then either an acquisition, merger or takeover would be preferable. According to different authors, liquidation emerges as a harvesting form when it cannot be prevented by the existing weak owners. Research implies that weak ownership stems from low levels of inside equity ownership, or high levels of outside equity ownership. It is important to note that a liquidation often occurs because the expectations of various stakeholders have not been met. The determinants of business failure are either the determinants that shape the conditions within the firm, the determinants associated with the firm's growth or determinants that reflect the firm's environment. The available research shows that problems in the market can emerge due to the concentration of customers, a firm's subcontracting status, a firm's export status and the presence of large competitors in the same region. An inadequate capital structure, high leverage, bad financial ratios, firm inefficiency and poor performance can also contribute to business failure and subsequent liquidation. Business failure also appears to be more likely for young and small firms. It can further be attributed to firms which have been overexpanding. A firm's past lower growth rate of employment is also found to be a factor that explains business failure. According to some results, unexpected macroeconomic and regulatory developments and uncertain legal and regulatory outcomes can also determine the probability of a business failure.

The literature review outlined in this chapter, therefore, supports the view that factors determining the selection of harvesting forms can indeed

be allocated into the four defined groups. In some cases, determinants are more specifically identified as in the case of IPOs or acquisitions, mergers and takeovers. In other cases, for example, with trade sales, they are defined in a broader sense. In the latter circumstances, the choice of this form of harvesting is selected due to investors' inability to select more attractive alternatives. Sometimes, as in the case of dividend payouts, however, harvesting is a consequence of investors' inability or resistance to exit the investment. That is why harvesting in the form of dividend payouts can be considered the "fundamental" form of harvesting. Some groups of determinants are rarely studied in the existing harvesting literature. This is especially true for determinants that shape the interests and decisions of owners and investors. This is most probably due to the unavailability of data. As a result, the number of determinants allocated to this group is fairly small and does not reflect the relative importance of this group of determinants. This clearly implies that the group of determinants shaping the interests and decisions of owners and investors calls for further research. Yet for other groups of harvesting forms the number of determinants included does reflect the relative importance of these groups in influencing the choice of various harvesting forms. Overall, the literature survey exploring the harvesting phenomenon confirms the thesis of this chapter that the form of harvesting reflects the interests of the investors, and that it also emerges from the relationship between the interests of the investors on one hand, and the firm's internal and external (environmental) characteristics that determine its growth possibilities on the other.

## REFERENCES

- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance* 23: 589–609.
- Altman, E. I. (1993). *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting and Avoiding Distress and Profiting from Bankruptcy*. John Wiley and Sons, New York.
- Ambrozini, V. (1998). *Exploring Techniques of Analysis and Evaluation in Strategic Management*. Prentice Hall Europe, London.
- Amit, R., Brander, J. and Zott, C. (1998). Why do Venture Capital Firms Exist? Theory and Canadian Evidence. *Journal of Business Venturing* 13: 441–466.
- Andrade, G. and Stafford, E. (2004). Investigating the Economic Role of Mergers. *Journal of Corporate Finance* 10: 1–36.
- Bandopadhyaya, A. and Jaggia, S. (2001). An Analysis of Second Time Around Bankruptcies Using a Split-Population Duration Model. *Journal of Empirical Finance* 8: 201–218.
- Bayless, M. and Chaplinski, S. (1996). Are there Windows of Opportunity for Seasoned Equity Issuance? *Journal of Finance* 51: 253–278.
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Empirical Research in Accounting: Selected Studies (Supplement to Journal of Accounting Research)* 4: 71–111.
- Becchetti, L. and Sierra, J. (2003). Bankruptcy Risk and Productive Efficiency in Manufacturing Firms. 27: 2099–2120.



- Berglöf, E. (1994). A Control Theory of Venture Capital Finance. *The Journal of Law, Economics and Organizations* 10: 247–267.
- Bethel, J. E., Liebeskind, J. P. and Opler, T. (1998). Block Share Purchases and Corporate Performance. *Journal of Finance* 53: 605–634.
- Bienz, C. (2004). *A Pecking Order of Venture Capital Exits—What Determines the Optimal Exit Channel for Venture Capital Backed Ventures?* Mimeo, <http://www.tu-dresden.de/vfs2004/Papers/Bienz.pdf>.
- Bjorvatn, K. (2004). Economic Integration and the Profitability of Cross-Border Mergers and Acquisitions. *European Economic Review* 48: 1211–1226.
- Black, B. S. and Gilson, R. J. (1998). Venture Capital and the Structure of Capital Markets: Banks versus Stock Markets. *Journal of Financial Economics* 47: 243–277.
- Bottazzi, L. and da Rin, M. (2003). *Financing Entrepreneurial Firms in Europe: Facts, Issues, and Research Agenda*. CESifo GmbH working paper, [http://ideas.repec.org/p/ces/ceswps/\\_958.html](http://ideas.repec.org/p/ces/ceswps/_958.html).
- Braganza, A. and Lambert, R. (2000). Strategic Integration: Developing a Process-Governance Framework. *Knowledge and Process Management* 7: 177–186.
- Brau, J., Francis W. and Kohers, N. (2003). The Choice between IPO versus Takeover: An Empirical Analysis. *Journal of Business* 76: 583–612.
- Brüderl, J. and Schüssler R. (1990). Organizational Mortality: the Liabilities of Newness and Adolescence. *Administrative Science Quarterly* 35: 530–547.
- Bygrave, W. D. and Timmons, J. (1992). *Venture Capital at the Crossroads*. Harvard Business School Press, Boston, MA.
- Carter, S. and Cannon, T. (1988). *Female Entrepreneurs*. Department of Employment Research Paper, University of Stirling, Scotland.
- Choe H., Masulis, R. W. and Nanda, V. (1993). Common Stock Offerings across the Business Cycle: Theory and Evidence. *Journal of Empirical Finance* 1: 3–33.
- Churchill, N. C. and Lewis, V. L. (1983). The Five Stages of Small Business Growth. *Harvard Business Review* 39: 30–50.
- Cochrane, J. H. (2001). *The Risk and Return of Venture Capital*. Working Paper No. 8066, NBER, Cambridge, MA.
- Cosh, A., Hughes, A. and Singh, A. (1984). *The Causes and Effects of Takeovers in the U.K.: an Empirical Investigation for the Late 1960s at the Microeconomic Level*. Economic Reprint No. 87. University of Cambridge, Department of Applied Economics, Cambridge.
- Cumming, D. J. (2002). *Contracts and Exits in Venture Capital Finance*. Working paper, University of Alberta, Canada, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=302695](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=302695).
- Cumming, D. J., Fleming, G. and Schwienbacher, A. (2003). *Liquidity of Exit Markets and Venture Capital Finance*. Presented at Third Workshop of the ECB-CFS Research Network on Capital Markets and Financial Integration in Europe, [http://www.eu-financial-system.org/November2003\\_Papers/A.Schwiebacher.pdf](http://www.eu-financial-system.org/November2003_Papers/A.Schwiebacher.pdf).
- Cumming, D. J., Fleming, G. and Schwienbacher, A. (2004). Legality and Venture Capital Exits. *Journal of Corporate Finance*, forthcoming.
- Cumming, D. J. and MacIntosh, J. G. (2001). Venture Capital Investment Duration in Canada and the United States. *Journal of Multinational Financial Management* 11: 445–463.
- Cumming, D. J. and MacIntosh, J. G. (2003a). A Cross-Country Comparison of Full and Partial Venture Capital Exits. *Journal of Banking and Finance* 27: 511–548.
- Cumming, D. J. and MacIntosh, J. G. (2003b). Venture Capital Exits in Canada and the United States. *University of Toronto Law Journal* 53: 101–200.
- Cumming, D. J. and MacIntosh, J. G. (2003c). Economic and Institutional Determinants of Venture Capital Investment Duration. *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth* 14: 125–159.

- Cumming, D. J. and MacIntosh, J. G. (2004). The Extent of Venture Capital Exits: Evidence from Canada and the United States. In *Venture Capital Contracting and the Valuation of High Technology Firms*, J. A. McCahery and L. Renneboog (Eds), pp. 339–371. Oxford University Press, Oxford.
- Das, S. R., Jagannathan, M. and Sarin, A. (2003). Private Equity Returns: An Empirical Examination of the Exit of Venture-Backed Companies. *Journal of Investment Management* 1: 1–26.
- David, F. R. (1999). *Strategic Management*, 7th issue. Prentice Hall, Upper Saddle River.
- Denis, D. J. and Denis, D. K. (1995). Causes of Financial Distress Following Leveraged Capitalizations. *Journal of Financial Economics* 37: 129–157.
- Denis, D. J. and Sarin, A. (1999). Ownership and Board Structures in Publicly Traded Corporations. *Journal of Financial Economics* 52: 187–223.
- Desroches, J. J.-Y. and Belletante, B. (1992). The Positive Impact of Going Public on Entrepreneurs and Their Firms: Evidence from Listing on the “Second Marche” in France. In *Frontiers of Entrepreneurship Research*, N. C. Churchill, J. Katz, B. Kirchoff, K. Vesper and W. Wetzel, Jr. (Eds), pp. 466–480. Babson College, Wellesley, MA.
- Desroches, J. J.-Y. and Jog, V. M. (1989). Evidence on the Impact of Going Public on the Entrepreneur and His Firm in Canada. In *Frontiers of Entrepreneurship Research*, R. H. Brockhaus, N. C. Churchill, J. Katz, B. Kirchoff, K. Vesper and W. Wetzel, Jr. (Eds), pp. 344–359. Babson College, Wellesley, MA.
- Dickerson, A. P., Gibson, H. D. and Tsakalotos, E. (2002). Takeover Risk and the Market for Corporate Control: The Experience of British Firms in the 1970s and 1980s. *International Journal of Industrial Organization* 20: 1167–1195.
- Didsbury, H. F. (1996). *Future Vision: Ideas, Insights, and Strategies*. World Future Society, Bethesda.
- Ding, D. K. and Sun, Q. (2001). Causes and Effects of Employee Stock Option Plans: Evidence from Singapore. *Pacific-Basin Finance Journal* 9: 563–599.
- Donoher, W. J. (2004). To File or Not To File? Systemic Incentives, Corporate Control, and the Bankruptcy Decision. *Journal of Management* 30: 239–262.
- Eisfeldt, A. and Rampini, A. (2005). *Capital Reallocation and Liquidity*. 7th Annual Texas Finance Festival Paper, <http://ssrn.com/abstract=384982>.
- Ekins, P., Hillman, M. and Hutchison, R. (1992). *Wealth Beyond Measure*. Gaia Books Limited, London.
- Erikson, T. and Sørheim, R. (2005). ‘Technology Angels’ and other Informal Investors. *Technovation* 25: 489–496.
- Faccio, M., Lang, L. H. P. and Young, L. (2001). Dividends and Expropriation. *American Economic Review* 91: 54–78.
- Fama, E. F. and French, K. R. (1992). The Cross-section of Expected Stock Returns. *Journal of Finance* 47: 427–465.
- Fleming, G. (2002). Exit Strategies in Venture Capital: Australian Evidence. *Venture Capital: An International Journal of Entrepreneurial Finance*, forthcoming.
- Foreman, R. D. (2003). A Logistic Analysis of Bankruptcy within the U.S. Local Telecommunications Industry. *Journal of Economics and Business* 55: 135–166.
- Franks, J. and Mayer, C. (1996). Hostile Takeovers and the Correction of Managerial Failure. *Journal of Financial Economics* 40: 163–181.
- Frydman, H., Altman, E. I. and Kao, D. L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *Journal of Finance* 40: 269–291.
- Gentry, J. A., Newbold, P. and Whitfort, D. T. (1987). Flow of Funds Components, Financial Ratios and Bankruptcy. *Journal of Business Finance and Accounting* 14: 595–606.

- Gladstone, D. and Gladstone, L. (2004). *Venture Capital Investing. The Complete Handbook for Investing in Private Businesses for Outstanding Profits*. Financial Times/Prentice Hall, New Jersey.
- Goergen, M. (1998). *Corporate Governance and Financial Performance: A Study of German and U.K. Initial Public Offerings*. Cheltenham, U.K.
- Gianaris, N. V. (1996). *Modern Capitalism*. Praeger Publishers, Westport.
- Giot, P. and Schwienbacher, A. (2003). IPOs, Trade Sales and Liquidation: Modeling Venture Capital Exits Using Survival Analysis, <http://econwp.wustl.edu/eps/fin/papers/0312/0312006.pdf> EconPapers (Online).
- Gleason, K. C., Rosenthal, L. and Wiggins, R. A. III (2004). Backing Into Being Public: an Exploratory Analysis of Reverse Takeovers. *Journal of Corporate Finance*, article in press.
- Gompers, P. (1996). Grandstanding in the Venture Capital Industry. *Journal of Financial Economics* 43: 133–156.
- Gompers, P. A. and Lerner, J. (1999a). What Drives Venture Capital Fundraising? Working Paper No. 6906, NBER, Cambridge, MA.
- Gompers, P. A. and Lerner, J. (1999b). Conflict of Interest in the Issuance of Public Securities: Evidence from Venture Capital. *Journal of Law and Economics* 42: 1–28.
- Gompers, P. A. and Lerner, J. (2004). *The Venture Capital Cycle, Second Edition*. The MIT Press, Cambridge, MA.
- Gort, M. (1969). An Economic Disturbance Theory of Mergers. *Quarterly Journal of Economics* 83: 623–642.
- Harford, J. (2005). What Drives Merger Waves? *Journal of Financial Economics*, article in press.
- Harhoff, D., Stahl, K. and Woywode, M. (1998). Legal Form, Growth and Exit of West German Firms: Empirical Results for Manufacturing, Construction, Trade and Service Industries. *Journal of Industrial Economics* 46: 453–488.
- Hasbrouck, J. (1985). The Characteristics of Takeover Targets: Q and Other Measures. *Journal of Banking and Finance* 9: 351–362.
- Hendrickson, L. (1988). *Bridging the Gap Between Organization Theory and the Practice of Managing Growth, An Organization Change Model of the Firm Development*. USASBE 3<sup>rd</sup> Annual Conference, Monterey, CA.
- Holmburg, S. R. (1991). Value Creation and Capture: Entrepreneurship Harvest and IPO Strategies. In *Frontiers of Entrepreneurship Research*, N. C. Churchill et al. (Eds), pp. 191–204. Babson College, Wellesley, MA.
- Hooper, A. and Potter, J. (1997). *The Business of Leadership*. Ashgate, Hants.
- Howorth, C., Westhead, P. and Wright, M. (2004). Buyouts, Information Asymmetry and the Family Managements Dyad. *Journal of Business Venturing* 19: 509–534.
- Huyghebaert, N. and Van Hulle, C. (2002). *Structuring the IPO: Empirical Evidence on the Primary and Secondary Portion*. Working paper. Katholieke Universiteit Leuven, Leuven.
- Hyatt, H. (1990). The Dark Side (of Going Public). *Inc.* 12: 46–56.
- Isaksson, A. (1998). *Venture Capital Exit Behaviour in Sweden*. Paper presented at the 10<sup>th</sup> Nordic Conference on Small Business research, Växjö University, Växjö, Sweden.
- Jeng, L. and Wells, P. (2000). The Determinants of Venture Capital Funding: Evidence Across Countries. *Journal of Corporate Finance* 6: 241–289.
- Jensen, M. C. (1986). Agency Cost of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review* 76: 323–329.
- Jensen, M. C. (1993). The Modern Industrial Revolution, Exit, and Control Systems. *Journal of Finance* 48: 831–880.
- Jensen, M. C. and Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* 3: 305–360.

- Jovanovic, B. and Rousseau, P. (2002). The Q-Theory of Mergers. Working paper No. 8740, NBER, Cambridge, MA.
- Kaplan, J. M. (2003). *Patterns of Entrepreneurship*. John Wiley and Sons Inc., Hoboken, NJ.
- Kensinger, J. W., Martin, J. D. and Petty, J. W. (2000). Harvesting Value from Entrepreneurial Success. *Journal of Applied Corporate Finance* 12: 81–93.
- King, N. (2002). *Exit Strategies*. Capstone Publishing, Oxford.
- Köke, J. (2002). Determinants of Acquisition and Failure: Evidence from Corporate Germany. *Structural Change and Economic Dynamics* 13: 457–484.
- La Porta, R., Lopez-De-Silanes, F., Shleifer, A. and Vishny, R. W. (2000). Agency Problems and Dividend Policies Around the World. *Journal of Finance* 55: 1–33.
- Lambrecht, B. M. (2004). The Timing and Terms of Mergers Motivated by Economies of Scale. *Journal of Financial Economics* 72: 41–62.
- Leland, H. and Pyle, D. (1977). Informational Asymmetries, Financial Structure and Financial Intermediation. *Journal of Finance* 32: 371–387.
- Lerner, J. (1994). Venture Capitalists and the Decision to Go Public. *Journal of Financial Economics* 35: 293–316.
- Levine, P. and Aaronovitch, S. (1981). The Financial Characteristics of Firms and Theories of Merger Activity. *Journal of Industrial Economics* 30: 149–172.
- Levinthal, D. A. (1991). Random Walks and Organizational Mortality. *Administrative Science Quarterly* 36: 397–420.
- Lin, T. H. and Smith, R. L. (1997). Insider Reputation and Selling Decisions: The Unwinding of Venture Capital Investments during Equity IPOs. *Journal of Corporate Finance* 4: 241–263.
- Low, S., Glorfeld, L., Hearth, D. and Rimbey, J. N. (2001). The Link between Bank Monitoring and Corporate Dividend Policy: The Case of Dividend Omissions. *Journal of Banking and Finance* 25: 2069–2087.
- Machlin, J. C., Cloe, H. and Miles, J. A. (1993). The Effects of Golden Parachutes on Takeover Activity. *Journal of Law and Economics* 36: 861–876.
- MacIntosh, J. G. (1997). Venture Capital Exits in Canada and in the United States. In *Financing Growth in Canada*, P. J. N. Halpern (Ed.), pp. 279–356. University of Calgary Press, Calgary.
- Maksimovic, V. and Phillips, G. (2001). The Market of Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains? *Journal of Finance* 56: 2019–2065.
- Mason, C. M. and Harrison, R. T. (2002). Is it worth it? The Rates of Return from Informal Venture Capital Investments. *Journal of Business Venturing* 17: 211–236.
- Matsunaga, S. R. (1995). The Effects of Financial Reporting Costs on the Use of Employee Stock Options. *Accounting Review* 70: 1–26.
- McGowan, J. (1971). International Comparisons of Merger Activity. *The Journal of Law and Economics* 14: 233–251.
- Meeks, G. (1977). *Disappointing Marriage: A Study of Gains from Merger*. Cambridge University Press, Cambridge.
- Meggison, W. C. and Weiss, K. A. (1991). Venture Capital Certification in Initial Public Offerings. *Journal of Finance* 46: 879–893.
- Meulbroek, L. K. (2000). *Does Risk Matter? Corporate Insider Transactions in Internet-Based Firms*. Unpublished Working Paper, Harvard Business School.
- Mikkelson, W. H., Partch, M. M. and Shah, K. (1997). Ownership and Operating Performance of Companies That Go Public. *Journal of Financial Economics* 44: 281–307.
- Mitchell, M. L. and Mulherin, J. H. (1996). The Impact of Industry Shocks on Takeover and Restructuring Activity. *Journal of Financial Economics* 41: 193–229.
- Mitton, T. (2004). Corporate Governance and Dividend Policy in Emerging Markets. *Emerging Markets Review* 5: 409–426.

- Morris, M. H. and Sexton, D. L. (1996). The Concept of Entrepreneurial Intensity: Implications for Company Performance. *Journal of Business Research* 36: 5–13.
- Mueller, D. C. (1969). A Theory of Conglomerate Mergers. *Quarterly Journal of Economics* 83: 643–659.
- Mulherin, J. H. and Boone, A. L. (2000). Comparing Acquisitions and Divestitures. *Journal of Corporate Finance* 6: 117–131.
- Napuk, K. (1993). *The Strategy-Led Business*. McGraw-Hill Book Company Europe, London.
- Opler, T. and Titman, S. (1994). Financial Distress and Corporate Performance. *Journal of Finance* 49: 1015–1040.
- Pagano, M., Panetta, F. and Zingales, L. (1998). Why do Companies go Public? An Empirical Analysis. *Journal of Finance* 53: 27–64.
- Pearce, J. A. and Robinson, R. B. (1991). *Strategic Management. Formulation, Implementation and Control*, 4th edition. Irwin, Boston, MA.
- Petty, J. W., Bygrave, W. D. and Shulman, J. M. (1994). Harvesting the Entrepreneurial Venture: A Time for Creating Value. *Journal of Applied Corporate Finance* 7: 48–58.
- Petty, J. W., Martin, J. D. and Kensing, J. W. (1999). *Harvesting Investments in Private Companies*. Financial Executives Research Foundation, Inc., Morristown, NJ.
- Petty, W. (1994). Harvesting. In *The Portable MBA in Interpreneurship*, W. D. Bygrave (Ed.), pp. 414–443. John Wiley & Sons, New York.
- Pompe, P. P. M. and Bilderbeek, J. (2004). The Prediction of Bankruptcy of Small- and Medium-Sized Industrial Firms. *Journal of Business Venturing*, article in press.
- Powell, R. and Yawson, A. (2005). Industry aspects of takeover and divestitures: Evidence from the U.K. *Journal of Banking and Finance*, article in press.
- Powell, R. G. (1997). Modeling Takeover Likelihood. *Journal of Business Finance and Accounting* 24: 1009–1030.
- Prasad, D. (1990). *An Asymmetric Information Approach to Forecasting Project Quality*. Ph.D. dissertation, University of Oklahoma, Norman, OK.
- Prasad, D. (1994). Is Underpricing Greater for Mixed Offerings as Compared to Pure Primary Offerings in the OTC Market. *Journal of Financial and Strategic Decisions* 7: 25–34.
- Prasad, D., Vozikis, G. S., Bruton, G. D. and Merikas, A. (1996). “Harvesting” Through Initial Public Offerings (IPOs): The Implication of Underpricing for the Small Firm. *Entrepreneurship: Theory and Practice* 20: 31–41.
- Rege, U. P. (1984). Accounting Ratios to Locate Take-over Targets. *Journal of Business Finance and Accounting* 11: 301–311.
- Relander, K., Syrjänen, A. P. and Miettinen, S. (1994). Analysis of the Trade Sale as a Venture Capital Exit Route. In *Realizing Investment Value*, W. D. Bygrave, M. G. Hay and J. B. Peeters (Eds), pp. 132–196. Pittman Publishing, London.
- Reuer, J. J. and Shen, J. (2003). The Extended Merger and Acquisition Process: Understanding the Role of IPOs in Corporate Strategy. *European Management Journal* 21: 192–198.
- Reuer, J. J. and Shen, J. (2004). Sequential Divestiture through Initial Public Offerings. *Journal of Economic Behavior and Organization* 54: 249–266.
- Rhodes-Kropf, M. and Viswanathan, S. (2004). Market Valuation and Merger Waves. *Journal of Finance* 59: 2685–2718.
- Rock, R. H. (1994). Economic Drivers of M&A. In *The Mergers and Acquisitions Handbook*, 2nd edition, M. L. Rock, R. H. Rock and M. Sikora (Eds), pp. 27–34. McGraw-Hill, New York.
- Rydqvist, K. and Högholm, K. (1995). Going Public in the 1980’s: Evidence from Sweden. *European Financial Management* 1: 287–315.
- Sahlman, W. A. and Hurlock, B. C. (1992). *Star Cablevision Group (A): Harvesting in a Bull Market*. Harvard Business School Case, September 17. Harvard Business School Publishing, Boston, MA.

- Sahlman, W. A. (1988). Aspects of Financial Contracting in Venture Capital. *Journal of Applied Corporate Finance* 1: 23–36.
- Sahlman, W. A. (1990). The Structure and Governance of Venture-Capital Organizations. *Journal of Financial Economics* 27: 473–521.
- Sahlman, W. A., Stevenson, H. H., Roberts, M. J. and Bhidé, A. (1999). *The Entrepreneurial Venture*. Harvard Business School Press, Boston, MA.
- Schwiebacher, A. (2005). *An Empirical Analysis of Venture Capital Exits in Europe and the United States*. EFA 2002 Berlin Meetings Discussion Paper, <http://ssrn.com/abstract=302001>.
- Sharplin, A. (1985). *Strategic Management*. McGraw Hill Book Company, New York.
- Shleifer, A. and Vishny, R. W. (1992). Liquidation Value and Debt Capacity: A Market Equilibrium Approach. *Journal of Finance* 47: 1343–1366.
- Shleifer, A. and Vishny, R. W. (1997). A Survey of Corporate Governance. *Journal of Finance* 52: 737–783.
- Shleifer, A. and Vishny, R. W. (2003). Stock Market Driven Acquisitions. *Journal of Financial Economics* 70: 295–311.
- Singh, A. (1971). *Takeovers: the Relevance to the Stock Market and the Theory of the Firm*. Cambridge University Press, Cambridge.
- Singh, A. (1975). Takeovers, Economic Natural Selection and the Theory of the Firm: Evidence from the Post-War U.K. Experience. *Economic Journal* 85: 497–515.
- Smith, D. G. (2001). *Control Over Exit in Venture Capital Relationships*. EFMA 2001 Lugano Meetings, <http://ssrn.com/abstract=272231>.
- Smith, L. R. and Smith, J. K. (2000). *Entrepreneurial Finance*. John Wiley & Sons, Inc., New York.
- Stewart, B. G. (1993). *Management and Employee Buy-outs in Central and Eastern Europe*. EBRD, London.
- Stuart, T. E., Hoang, H. and Hybels, R. C. (1999). Interorganizational Endorsements and the Performance of Entrepreneurial Ventures. *Administrative Science Quarterly* 44: 315–349.
- Tajnikar, M. and Brščič, B. (2002). Characteristics of High Growth Firms in Late Transition Circumstances—The Case of Slovenia. In *Dynamic Entrepreneurship for New Economy, 3rd International Conference Proceedings*, P. Davidsson, P. Barlett, W. Žižek and J. Vadnjal (Eds), pp. 24–34. Visoka strokovna šola za podjetništvo, Ljubljana, Slovenia.
- Tajnikar, M. and Došenovič, P. (2003). *Harvesting in Slovene Fast-Growing Companies*. Paper presented at 2003 International Entrepreneurship & New Venture Creation Conference, Durham Business School, Durham, U.K.
- Thompson, A. A. and Strickland, A. J. (1998). *Crafting and Implementing Strategy*, 10th edition. Irwin McGraw-Hill, Boston, MA.
- Thompson, S. (1997). Takeover Activity among Financial Mutuals: An Analysis of Target Characteristics. *Journal of Banking and Finance* 21: 37–53.
- Timmons, J. A. (1990). *Planning and Financing the New Venture*. Brick House Publishing Company, Acton, MA.
- Timmons, J. A. (1999). *New Venture Creation*, 5th edition. McGraw-Hill Book Co., Singapore.
- Upton, N. and Petty, W. (1998). *Funding Options for Transferring the Family Health Firm: A Comparative Analysis*. Paper presented at the 18<sup>th</sup> Babson Kaufman Entrepreneurship Research Conference, Babson College, Wellesley, MA.
- Van Osnabrugge, M. and Robinson, R. J. (2000). *Angel Investing*. Jossey-Bass Inc., San Francisco, CA.
- Wall, J. and Smith, J. (1998). *Better Exits*. Technical Report. European Venture Capital Association, Brussels.
- Wickham, P. A. (1998). *Strategic Entrepreneurship*. Pitman Publishing, London.

- Wright, M., Robbie, K., Romanet, Y., Thompson, S., Joachimsson, R., Bruining, J. and Herst, A. (1993). Harvesting and the Longevity of Management Buy-outs and Buy-ins: A Four-Country Study. *Entrepreneurship Theory and Practice* 18: 90–106.
- Yermack, D. (1995). Do Corporations Award CEO Stock Options Effectively? *Journal of financial Economics* 39: 237–269.
- Zingales, L. (1995). Insider Ownership and the Decision to Go Public. *Review of Economic Studies* 62: 425–448.
- Zingales, L. (1998). Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry. *Journal of Finance* 53: 905–938.
- Zmijewsky, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* 22: 59–82.

# Index

- Ability 7, 210, 325, 338, 340, 342, 346–347, 357, 373, 492–493, 509, 512, 514, 521, 523, 537
- Age 21, 25–26, 32, 65, 118, 207, 211, 263–264, 354–355, 371, 418, 453, 467, 485–486, 489–492, 496, 500, 551, 568
- Bank loans *See* Finance, Debt
- Bankruptcy and liquidation 10, 528, 540, 545–546, 566–569
- Borrowing constraints *See* Credit constraints
- Business angels *See* Finance, Informal
- Business plans 20, 32, 84, 87, 235, 240, 498
- Buybacks 540, 546, 556–560, *see also* Management buy-outs/buy-ins
- Collateral 5, 187, 195, 198–204, 215–216, 219–221, 301, 304–307, 309–310, 314, 321–322, 327, 516, 519, 521, 560
- Credit constraints 5–7, 10, 145–146, 165, 187–188, 204–221, 231, 340, 454–455, 462, 515–530
- Culture 18, 19, 56, 107–108
- Decomposition 114–115, 118–123, 124–127, 439, 460–465, 468–469
- Education (schooling) *See* Human Capital
- Effort 6, 30, 232–233, 235–253, 338–340, 349–355, 357, 538, 553
- Entrepreneurship capital 7, 338, 342, 347–349, 357, 537
- Entrepreneurship education 3–4, 7, 79–100, 347, 357, 527, 528–529
- Entry 4–5, 25, 43, 91, 108–109, 137–156, 161–179, 230, 233, 240–243, 247–251, 340, 353–355, 357, 384, 410, 418, 438–439, 444–456, 460–464, 466–471, 480, 484, 507, 517, 520, 529, 573
- Ethnicity *See* Minority Entrepreneurship
- Exit and Survival 3–5, 7, 9–10, 15, 25, 27–28, 39, 66, 86, 161–179, 189, 203, 208, 231, 247–248, 253, 340, 355–356, 369, 382, 438–439, 444–449, 456–470, 479–503, 507, 511, 514, 519–520, 523, 525, 530, 535–537, 539–542, 545–546, 566
- Experience 26–28, 31–32, 42, 66–67, 87, 139, 164–165, 203, 262, 279–280, 341, 371, 373, 486–488, 493–495, 500, 508, 518, 543
- Failure *See* Exit
- Female entrepreneurship 7–9, 96, 325, 354, 357, 371, 440–442, 445, 455–456, 467–468, 496
- Finance 5–7, 9, 26, 32, 84, 139, 165, 187–223, 227–254, 259–292, 301–330, 347, 350, 375, 419, 422–423, 438, 465, 483, 507–530, 535–575
- Debt 5, 187–223, 227, 244, 249, 271, 279, 283, 422, 516, 529–530, 545, 548, 552–553, 560–561, 567, 569, 573
- Equity (venture capital) 5–6, 79–80, 146, 218–219, 227–254, 267, 269, 272, 278, 283–284, 287, 535–537, 539–540, 542, 545–547, 550–551, 553–554, 556
- Informal 5–6, 259–292, 540–541, 545, 557
- Micro-finance 5, 7, 301–330



- Gender 21, 23–24, 26–27, 30, 32, 263, 354–355, 371, 437–471, 479, 486–487, 491, 496, 499–500
- Geographic location 4, 28, 137–156, 171–173, 269–270, 276, 281, 315, 348, 368–369, 371, 377–378, 453, 464–465, 467, 569, 574
- Global Entrepreneurship Monitor (GEM) 16–17, 19, 21–23, 26, 33, 63, 162–163, 259
- Grameen Bank 301, 306, 309, 316, 321, 330, 401
- Growth 7–8, 9, 16, 39, 41, 43, 49, 63–64, 86, 167–169, 189–190, 212, 228–232, 255, 278, 344–345, 348, 355–356, 361–390, 405, 423, 425, 429, 479–503, 507, 508, 511, 519, 543–544, 549, 551–552, 561, 563, 568
- Harvesting 9, 10, 273–274, 282–284, 535–575  
 Forms of harvesting 535, 540  
 Elements of 538  
 Timing 539, 551
- Human capital 9–10, 19, 23, 26, 30, 32, 55, 64, 84, 113, 118, 130, 138, 140, 144–145, 151–152, 155–156, 162, 164–165, 172, 195, 207–211, 264, 304, 325, 348, 372–373, 418, 422, 437–471, 479, 482–483, 485–488, 491, 493, 498, 500, 502, 507–530
- Immigration 4, 107–133, 446, 448–450, 453–457, 460, 462–466, 468, 479
- Industry life cycle 2, 169–170, 173
- Initial public offering (IPO) 10, 149, 155, 230–231, 238, 247, 263, 284, 540, 545–546, 549–557, 565, 569
- Innovation 2, 7, 18–19, 44–45, 49, 53–54, 60, 66, 138, 140, 142–143, 145, 147–148, 155, 161, 163, 170–171, 175, 228, 232, 252, 273, 338, 340, 342–345, 348, 357, 369, 372, 402, 412, 419–420, 507
- International entrepreneurship 377–378
- Intrapreneurship 2, 15, 65–67, 69, 162, 376
- Jack-of-all-trades hypothesis 27, 164–165, 511
- Knowledge spillovers 4, 7, 137–156, 345, 347–348, 354, 357
- Latent entrepreneurs 15, 438, 456, 516
- Liquidation *See* Bankruptcy and liquidation
- Management buy-outs/buy-ins 10, 540, 545–546, 556–561
- Mergers, acquisitions 10, 362, 364, 367, 370, 375–377, 385, 388, 540, 545–546, 550, 552, 562–566
- Minority entrepreneurship 8–9, 26, 96, 108, 371, 418–419, 437–471, 479, 496–497, 500–501, 507
- Mutual guarantee schemes 217–218
- Nascent entrepreneurs 2–3, 15–33
- Non-distribution constraint 404–406, 410, 419
- Not for profit enterprises *See* Social entrepreneurship
- Occupational choice 24–25, 138–139, 205–207, 235, 240–243, 251, 354
- Opportunities 3, 7, 39, 46–47, 81, 83, 139, 161, 163, 167, 170, 178, 275–277, 289, 345, 364, 370, 373, 402, 419–420, 424, 428, 438, 482, 493–494, 497–500, 501, 544, 549, 563
- Opportunity versus necessity entrepreneurship 19, 63
- Panel Study of Entrepreneurial Dynamics (PSED) 15–16, 20, 23, 30, 33, 139–140
- Production and production functions 7, 142–143, 337–357
- Psychology of entrepreneurs (including risk attitudes) 2, 7, 83, 163–164, 213–214, 264, 347, 357, 374, 417, 508
- Public policy 6, 9, 10, 57–59, 79, 107, 110–113, 121–122, 128–130, 173, 212, 219–220, 227–240, 242–254, 270–273, 288–289, 314, 322–323, 328–329, 337, 351, 353, 355, 357, 374–375, 402, 437–438, 479, 482, 483–484, 489, 491, 496, 500–501, 507, 527–529
- Regions *See* Geographical Location
- Regional Entrepreneurship Monitor (REM) 27–29
- Relationship banking 215–217

- Resource-based view (RBV) 40, 49, 54–55, 69, 156, 374
- Risk *See* Uncertainty
- Secondary sale 540, 545–546, 555
- Self-employment 29–30, 107–133, 189, 205–208, 325, 341, 351–353, 371, 437–471, 480, 484–496, 501
- Small firms 28–29, 141, 163, 170, 188–195, 198, 216, 304, 341, 346–347, 361–390, 479–483, 485, 489–493, 499, 501–502, 549, 551–552, 560, 565, 567, 568, 573–574
- Social capital 32, 55, 139–140, 162, 165, 275, 310, 402, 426, 427, 429, 524–525
- Social entrepreneurship (including not for profit ventures) 8, 58–59, 70–71, 355, 401–430
- Strategy of entrepreneurs 2–3, 7, 39–71, 146, 156, 175–177, 369, 372, 374, 429, 493–494, 497–499, 501, 508, 537–539, 541–545, 549, 553, 560–561, 572
- Survival *See* Exit
- Taxes and subsidies 6, 228–231, 234–254, 271–272, 351, 353, 356, 405, 411, 416, 418–419, 424, 528, 538, 561, 573
- Technology *See* Innovation and technology
- Theory of the firm 2, 364
- Trade sale 284, 540, 545–546, 556–562, 567
- Uncertainty and Risk 63, 94, 274, 278, 282–283, 305–306, 310–311, 429, 543, 549, 563–564
- Unemployment 172, 371, 406, 419, 421, 479, 482, 483–488, 490–491, 500, 508
- Universities 24, 79–80, 83–85, 88–93, 98, 137–138, 143–156, 413
- Venture capital *See* Finance, Equity (venture capital)

