PALGRAVE READERS IN ECONOMICS

Global Labour in Distress, Volume I

Globalization, Technology and Labour Resilience

Edited by
Pedro Goulart · Raul Ramos
Gianluca Ferrittu

palgrave
macmillan

Palgrave Readers in Economics

This series brings together previously published papers by leading scholars to create authoritative and timely collections that contribute to economic debate across a range of topics. These volumes are aimed at graduate level student and beyond to provide introductions to, and coverage of, key areas across the discipline.

Pedro Goulart • Raul Ramos
Gianluca Ferrittu
Editors

Global Labour in Distress, Volume I

Globalization, Technology and Labour Resilience

palgrave
macmillan

Editors
Pedro Goulart
CAPP, Institute of Social and Political Sciences
Universidade de Lisboa
Lisbon, Portugal

Raul Ramos
Department of Econometrics, Statistics and Applied Economics, AQR-IREA
University of Barcelona
Barcelona, Spain

Gianluca Ferrittu
Lisbon School of Economics and Management - ISEG
Universidade de Lisboa
Lisbon, Portugal

ISSN 2662-6454 ISSN 2662-6462 (electronic)
Palgrave Readers in Economics
ISBN 978-3-030-89257-9 ISBN 978-3-030-89258-6 (eBook)
https://doi.org/10.1007/978-3-030-89258-6

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Article Note

This first Volume, "Globalisation, Technology and Labour Resilience" collects 13 new contributions and 15 previously published journal articles in selected Springer Journals. Reprints in this volume are included as different chapters with the permission of the author and the publisher. Further details on previously published chapters are provided below.

Chapter No.	*Original Citation*
3	Mayda, A. M. (2010). International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. *Journal of Population Economics, 23*, 1249–1274. https://doi.org/10.1007/s00148-009-0251-x
6	Saxenian, A. (2005). From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China. *Studies in Comparative International Development, 40*, 35–61. https://doi.org/10.1007/BF02686293
7	Dluhosch, B., & Horgos, D. (2019). International Competition Intensified: Job Satisfaction Sacrificed? *Social Indicators Research, 143*, 479–504. https://doi.org/10.1007/s11205-018-1982-4
8	Magda, I., & Sałach, K. (2021). Gender Pay Gaps in Domestic and Foreign-Owned Firms. *Empirical Economics, 61*, 2237–2263. https://doi.org/10.1007/s00181-020-01950-z
12	Galor, O., & Stark, O. (1991). The Impact of Differences in the Levels of Technology on International Labor Migration. *Journal of Population Economics, 4*, 1–12. https://doi.org/10.1007/BF00160365

(*continued*)

(continued)

Chapter No.	*Original Citation*
13	Rudra, N. (2005). Are Workers in the Developing World Winners or Losers in the Current Era of Globalization? *Studies in Comparative International Development, 40*, 29–64. https://doi.org/10.1007/BF02686298
15	Kancs, d., & Siliverstovs, B. (2020). Employment Effect of Innovation. *Empirical Economics, 59*, 1373–1391. https://doi.org/10.1007/s00181-019-01712-6
16	Abbott, P., Tarp, F., & Wu, C. (2017). Structural Transformation, Biased Technological Change and Employment in Vietnam. *The European Journal of Development Research, 29*, 54–72.
17	Berner, E., Gomez, G., & Knorringa, P. (2012). 'Helping a Large Number of People Become a Little Less Poor': The Logic of Survival Entrepreneurs. *The European Journal of Development Research, 24*, 382–396. https://doi.org/10.1057/ejdr.2011.61
18	Peša, I. (2018). The Developmental Potential of Frugal Innovation among Mobile Money Agents in Kitwe, Zambia. *The European Journal of Development Research, 30*, 49–65. https://doi.org/10.1057/s41287-017-0114-3
22	Smith, J., Thomas, D., Frankenberg, E., Beegle, K., & Teruel, G. (2002). Wages, Employment and Economic Shocks: Evidence from Indonesia. *Journal of Population Economics, 15*, 161–193. https://doi.org/10.1007/PL00003837
23	Rafferty, A., Rees, J., Sensier, M., & Harding, A. (2013). Growth and Recession: Underemployment and the Labour Market in the North of England. *Applied Spatial Analysis and Policy, 6*, 143–163. https://doi.org/10.1007/s12061-013-9089-4
24	Arpino, B., Pronzato, C. D., & Tavares, L. P. (2014). The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach. *European Journal of Population, 30*, 369–390. https://doi.org/10.1007/s10680-014-9319-8
26	López-Calva, L. F., & Ortiz-Juarez, E. (2014). A Vulnerability Approach to the Definition of the Middle Class. *The Journal of Economic Inequality, 12*, 23–47. https://doi.org/10.1007/s10888-012-9240-5
28	Finch, C., Emrich, C. T., & Cutter, S. L. (2010). Disaster Disparities and Differential Recovery in New Orleans. *Population and Environment, 31*, 179–202. https://doi.org/10.1007/s11111-009-0099-8

Preface

The end to the Cold War and a unipolar world coincided with the retrenchment of the state and a move towards a more market-based economy. Since then, what developments occurred in the world of labour? Globalization and technology provoked a major change in the economic production, while schooling has been expanded and democratized throughout the globe, with developing countries at different stages now educating most of their youths. But what about labour conditions and the inequality of earnings? How resilient has been labour to adapt to these changes? How did labour institutions and policies evolve?

Over the last 30 years, the power of labour showed, at best, contradictory signs or even became considerably frailer. Following the earlier elections of Thatcher and Reagan and contributing to the declining formal labour protection, developed countries experienced falling union rates and social concertation practices. Segmented labour markets led often to polarization of labour earnings and conditions. Progresses in living standards and in different spheres in human development, noticeably in Asia and particularly China, made starker the inequality in the Global South. Since the 1990s, there has also been a deceleration in the liberalization of migration policies that predominated since WWII, while international migration has remained remarkably stable contributing to the higher complexity and diversity of labour markets. What would be coined as the "Washington Consensus" neglected labour conditions in developing countries.

Labour institutions were thus challenged throughout the period. Informal labour, female and youth and child labour received increasing attention and resources. However, recurrent crises, high unemployment,

youth unemployment and underemployment, low wages and, in general, falling wage shares have been motives for discontentment and even upheaval of many. Intersectionality often comes to make more severe the situation of some. A more globalized and interconnected world also made inequalities more visible, discomforting and conflictive. In sum, the age post-1991 was marked by *global labour in distress*.

This two-volume book aims to frame these issues by composing a 30-year perspective, which allows for an uncommon depth of analysis. By doing so, *Global Labour in Distress* collects selected high-influential journal articles on labour issues around the globe since 1991 and compiles them with brief unpublished masterpieces defined by highly recognized experts, providing a complementary view from today's perspective. Building on the editors' expertise on economics, demography and development studies, the journal articles were selected from a pool of highly reputed Springer journals among the academic community. Their varied disciplinary approaches provide a multidisciplinary perspective to labour issues.

The selection was based on a pool of 15,047 journal articles of 1474 issues of 11 scientific journals, as identified in Table 1. For the selection of the articles, the editorial team searched for topics related to labour markets and globalization; jobs and technological change; labour agency and resilience; labour earnings and inequality; (in)decent work; continuity and change of labour institutions; gender, in a crosscutting manner. In a first stage, the editors chose circa 60 papers, to, in a second stage, reduce it to

Table 1 Springer journals included in the selection

Areas	*Journals*	*Number of articles*
Economics	*Empirical Economics*	12
	Journal of Family and Economic Issues	
	Journal of Economic Inequality	
	Journal of Population Economics	
Development	*Applied Spatial Analysis and Policy*	11
	European Journal of Development Research	
	Studies in Comparative International Development	
Demography and other social sciences	*European Journal of Population*	7
	Journal of Labor Research	
	Population and Environment	
	Social Indicators Research	

the final list. In the selection of the papers, we privileged the quality of the papers, the variety of topics and the diversity of affiliations/institutions. The final volumes gather "[a]n amazing line up of great authors," in the words of one contributor. Fifty-six chapters, from 91 authors affiliated to institutions from 22 countries, covering different regions of the world. After the start of the project, one of the authors was actually awarded the Nobel Prize in Economic Sciences, we leave it to the reader to identify who. The geographical variety of each section is fundamental given the importance of studying the context. The two volumes feature complementary topics on labour issues, but sometimes with opposing views.

Each volume is structured in three main sections and analyses well-defined but also complementary topics on labour issues. All sections include an introduction by the editors followed by a compilation of selected articles' selection and masterpieces. In the first volume, *Globalisation, Technology and Labour Resilience*, the first section is about developments that occurred in the world of labour related to the integration of labour markets and globalization, bridging international migration drivers and policy and the level of internationalization of production. The second section analyses technological change and innovation, discussing structural transformation and frugal innovation, employment and jobs adaptation, multinationals and survival entrepreneurship. Finally, the third section discusses the change in labour agency and resilience concerning major changes in international and national institutional landscapes, discussing informalization of labour and underemployment, the politics of workplace wellbeing and the effects of crises and their recovery.

The second volume *Earnings, (In)decent work and Institutions* follows a similar structure. It is also structured into three main sections. The first one addresses earnings and inequality, bridging trade globalization and COVID-19 pandemic effects, the geography of poverty, horizontal inequality and inequality of opportunity, unions' impact on wages and the gender gap. The second one focuses on the analysis of recent trends in decent work, discussing labour standards, unemployment and minimum wages and gender issues and work-family balance. Last, the third section discusses the role of labour market policies and its interactions with institutions, and it combines pieces on growth and labour standards, social protection policies and policy tools.

Lisbon, Portugal — Pedro Goulart
Barcelona, Spain — Raul Ramos
Lisbon, Portugal — Gianluca Ferrittu

Acknowledgements

As always, this endeavour has only been possible by the contribution of many.

First and foremost, we need to thank the 88 contributors that have produced so many enriching views on labour and who so selflessly agreed to join this effort. A special word goes to the very few that agreed but for unforeseen reasons were not able to join in the end.

The community around the selected journals, the editors and reviewers are also to be praised as our selection benefited from their constant and careful mostly unpaid work. Science, although imperfect and incomplete, has allowed us a long way in understanding social phenomena.

We would like to thank Spencer Henson, Natalia Lorenzoni and Wendy Olsen for helpful comments on an earlier phase and to Catarina Grilo.

Wyndham Hacket Pain and Steve Fashions at Palgrave/Springer have been extremely diligent and competent editors in supporting the volumes.

We also gratefully acknowledge the financial support received from the Portuguese Ministry of Science and Technology and the Foundation for Science and Technology through the projects UIDB/00713/2020 (to Pedro Goulart) and 2020.05445.BD (to Gianluca Ferrittu), and, the Spanish Ministry of Science and Innovation and the State Research Agency (MCIN/AEI—10.13039/501100011033) through the project PID2020-118800GB-I00 (to Raul Ramos).

On a personal note, we also need to thank our families who have abdicated part of our precious time together. At this moment, we are thinking particularly on:

To my father, uncle, aunt and Zé who are dearly missed, to my mother and her strength, to C. for 21 wonderful years and to A. who wanted to help by drawing pictures to illustrate the book (Pedro Goulart).

To my wife and lovely daughters, and to my parents from whom I learned first-hand experience on some labour-related issues like child work or rural to urban migration, but also resilience. Thanks for everything (Raul Ramos).

To my parents, M. and P., my sisters, G. and S., and my lovely G., for their unconditional support and love. To my grandad L., who is always with me (Gianluca Ferrittu).

Contents

Notes on Contributors

Philip Abbott is Professor Emeritus of Agricultural Economics at Purdue University, West Lafayette, IN, USA. He conducted research on both international trade and international agricultural development. He has consulted for several domestic and foreign government agencies, the United Nations Food and Agriculture Organization (FAO), the OECD, the World Bank, commissions on food policy issues and private agencies. He was on the editorial boards of the *American Journal of Agricultural Economics* and the *Journal of Development Economics.* He served on the steering committees of several national research organizations.

Vito Amendolagine is a Tenure Track Assistant Professor of Economics at the University of Foggia (Italy). His research is focused on international trade and development economics, with particular attention to foreign direct investments in Sub-Sahara African economies and technological investments from emerging markets. He has got an MSc in Economics and Econometrics from the University of Essex, a PhD in Economics from the University of Bari and a PhD in Economics from the University of Glasgow. He has working experience with UNIDO, London School of Economics, Aalborg University, University of Pavia and Bari.

Bruno Arpino is an associate professor at the University of Florence (Italy). Previously he was an associate professor at the Universitat Pompeu Fabra (UPF, Spain) and co-director of the Research and Expertise Center on Survey Methodology (RECSM, UPF). His main research interests are in the areas of social demography and applied statistics. He is particularly interested in studying intergenerational relationships, ageing and health,

and fertility. He coordinated an international project titled "Care, Retirement & Wellbeing of Older People Across Different Welfare Regimes" (CREW). His publications appear in several international journals such as *Demography*, *Journal of Marriage and Family*, *The Journals of Gerontology: Series B*, *European Sociological Review*. For more info see his personal webpage: https://sites.google.com/site/brunoarpino/

Laura Bartolini is a researcher and data analyst under the Displacement Tracking Matrix (DTM) programme of IOM since 2016. Within DTM, she supports the collection and harmonization of data on mixed migration flows across the Mediterranean, Western African Atlantic and Western Balkan routes to Europe, including primary data collection of surveys with people on the move. She works on research and data collection activities on migrants in vulnerable situations, including victims of exploitation, abuse and trafficking, and unaccompanied migrant children. Previously, she was research associate at the European University Institute (2011–2016) and collaborated, among others, with OECD, the World Bank, FIERI.

Kathleen Beegle is Research Manager and Lead Economist in the World Bank's Development Research Group. Her research experience includes the study of poverty, labour, economic shocks and methodological studies on household survey data collection in developing countries. She has expertise in the design and implementation of household surveys and their use for poverty and policy analysis. She received her PhD in Economics from Michigan State University and completed a post-doctoral fellowship at RAND.

Peter A. G. van Bergeijk (1959) is Full Professor of International Economics and Macroeconomics at the International Institute of Social Studies of Erasmus University, the Netherlands. His most recent publications include the monographs *Pandemic Economics* (2021) and *Deglobalization 2.0: Trade and openness During the Great Depression and the Great Recession* (2019) and the *Research Handbook on Economic Diplomacy: Bilateral Relations in a Context of Geopolitical Change* (2018, co-edited with Selwyn Moons). Peter is editor of the *Research Handbook on Economic Sanctions.*

Erhard Berner is a retired Associate Professor of Development Sociology at the Institute of Social Studies, The Hague, The Netherlands, and a visiting scholar at the University of Witwatersrand, Johannesburg, South

Africa. He has done extensive research on urban poverty and community responses in several countries in Africa, South and Southeast Asia and published a book and numerous articles on the subject. He has also served as a consultant to UN-Habitat, international and Dutch NGOs and government institutions in the fields of urban poverty, housing, basic services and small enterprise promotion. Current research focuses on informal housing, microenterprises and poverty alleviation.

Elizabeth Cotton is a writer and educator in the field of mental health at work and an academic at Cardiff School of Management at Cardiff Metropolitan University. Her background is as an international educator working for global union federations and as an honorary psychotherapist in the National Health Service. She is an editor-in-chief of AJG4 journal *Work Employment & Society* (WES) and is a British Sociological Association Trustee and Publications' Director. She is a member of the UK's Health & Safety Executive's Science and Engineering Evidence and Assurance Committee (SEEAC). She is founding director of www.survivingwork.org and her current research on digitalization and mental health can be accessed at www.thefutureoftherapy.org.

Susan Cutter is a Carolina Distinguished Professor of Geography at the University of South Carolina where she directs the Hazards and Vulnerability Research Institute. She received her BA from California State University, East Bay, and her MA and PhD (1976) from the University of Chicago. Her primary research interests are in the area of disaster vulnerability/resilience science and how vulnerability and resilience are measured, monitored and assessed. She has authored or edited 14 books, the most recent book is *Hurricane Katrina and the Forgotten Coast of Mississippi*, more than 150 peer-reviewed articles and book chapters. Cutter has mentored more than 50 master's and doctoral students.

Keshab Das is a professor at the Gujarat Institute of Development Research, Ahmedabad, India. He holds MPhil (Applied Economics) and PhD (Economics) degrees from the Jawaharlal Nehru University, New Delhi (through the Centre for Development Studies, Trivandrum). He is a recipient of the VKRV Rao Prize in Social Sciences (Economics). He is the Executive Editor of the *Orissa Economic Journal*. His research focuses on issues in local and regional development; industrialization strategies; informal sector; MSMEs, industrial clusters and globalization; innovation; labour; basic infrastructure; and politics of development.

Barbara Dluhosch is Full Professor of International Economics at the Helmut Schmidt University in Hamburg. Her previous positions include fixed-term Professor of Macroeconomics at the Ludwig-Maximilians-University in Munich and visiting positions at Stanford University and the Bank of Spain in Madrid. Her research interest is primarily focused on the fields of international economics (trade and finance), in particular, the globalization of production, information and communication technologies and their impact on labour markets and subjective well-being as well as geopolitical issues related to international trade.

Werner Eichhorst (1969) studied sociology, political science, psychology and public policy and administration at the universities of Tuebingen and Konstanz where he graduated in 1995. From 1996 to 1999 he was doctoral and post-doctoral fellow at the Max Planck Institute for the Study of Societies in Cologne. In fall 1998 he received his doctoral degree from the University of Konstanz. From 1999 to 2004 he was project director at the Bertelsmann Foundation, a private think tank in Germany. After working with the Institute for Employment Research, IAB, from 2004 to 2005, he joined IZA where he works as coordinator of Labor Market and Social Policy in Europa since January 2017 and Team Leader Research since August 2019. Since November 2017 he is honorary professor at Bremen University. His main research area is the comparative analysis of labour market institutions and performance as well as the political economy of labour market reform strategies. He also specializes in different aspects of the future of labour.

Christopher Emrich is an Endowed Associate Professor of Environmental Science and Public Administration within the School of Public Administration and a founding member of the newly formed National Center for Integrated Coastal Research at University of Central Florida (UCF Coastal). His research/practical service includes applying geospatial technologies to emergency management planning and practice, long-term disaster recovery and the intersection of social vulnerability and community resilience in the face of catastrophe. From 2004–2007, he provided geospatial support for response and long-term recovery to the state of Florida and has since been actively involved in identifying long-term recovery trends and drivers in disaster-stricken areas nationwide.

Alessandra Faggian is Professor of Applied Economics, Director of Social Sciences and Deputy Rector at the Gran Sasso Science Institute,

L'Aquila, Italy. She is co-editor of *Journal of Regional Science* and previous editor of *Papers in Regional Science*. She is vice president of the Italian Economics Society (SIE) and has been president of the North American Regional Science Council (NARSC). Alessandra is the 2007 recipient of the Moss Madden Memorial Medal by the Regional Science Association International: Irish and British section (RSAIBIS), the 2015 recipient of the Geoffrey Hewings Award by NARSC, the 2020 European Regional Science Association (ERSA) Prize winner and was elected Fellow of the Regional Science Association International (RSAI) in 2021.

Gianluca Ferrittu is a doctoral researcher at the Lisbon School of Economics and Management - ISEG (Universidade de Lisboa), and he holds a master's degree from the University of Pavia in Economics, Politics and International Institutions. He has obtained a Foundation for Science and Technology studentship grant (economics panel, 2020), and a Giovanni Manera Fellowship (University of Pavia, 2018). His research interests lay particularly on labour markets and public policy, with a focus on child labour and decent work. Previously, he has worked in the research unit of the United Nations Migration Agency (IOM Italy) and has been hosted at the Centre for Public Administration and Public Policies (ULisboa), and at the Myanmar Centre for Responsible Business (Yangon).

Christina Finch currently works for the Federal Emergency Management Agency (FEMA) where she has been working in emergency management, including deployments during the 2004 Hurricane Season and Hurricane Katrina. Prior to FEMA, Christina served as the Senior Risk and Vulnerability Assessment Analyst at the Pacific Disaster Center in Maui, HI. Finally, she has served in multiple roles at the Hazards and Vulnerability Research Institute at the University of South Carolina, and she is currently an adjunct lecturer at Georgetown University where she teaches an online course related to the Socio-Cultural Dimensions of Emergency Management in the School of Continuing Studies Emergency & Disaster Management Department.

Elizabeth Frankenberg is the Cary C. Boshamer Distinguished Professor of Sociology at the University of North Carolina at Chapel Hill and Director of the Carolina Population Center. She co-directed waves 1, 2 and 2+ of the Indonesia Family Life Survey. Her research focuses on individual and family responses to change and the role of community, broadly construed, in behaviours and outcomes across the life course. Cross-

cutting themes inherent in her research are health status as a critical dimension of well-being and the close integration of methods and data. She holds a PhD in Demography from the University of Pennsylvania.

Oded Galor is the founder of Unified Growth Theory. He has contributed to the understanding of the process of development over the entire course of human history and the role of deep-rooted factors in the transition from stagnation to growth and in the emergence of the vast inequality across the globe. Moreover, he has pioneered the exploration of the impact of evolutionary processes, population diversity and inequality on the process of development over most of human existence. His research has redirected research in the field of economic growth to the exploration of the long shadow of history and to the role of biogeographical forces in comparative economic development. It has spawned the influential literatures studying the impact of inequality on the process of development, the interaction between human adaptation and economic development, the transition from stagnation to growth and the impact of human diversity on comparative economic development. Oded Galor was awarded Doctor Honoris Causa from UC Louvain and from Poznań University of Economics & Business. He is an elected Foreign Member of Academia Europaea (honoris causa) and an Elected Fellow of the Econometric Society. He has co-directed the NBER research group on Income Distribution and Macroeconomics and a Research Fellow of CEPR, GLO and IZA, a research associate of the NBER and CESifo, a Sackler Fellow at Tel-Aviv University, a Fellow of the Economics Department at the Hebrew University. Furthermore, he is the editor-in-chief of the *Journal of Economic Growth*, editor of the *Journal of Population Economics* and co-editor of *Macroeconomic Dynamics*.

Georgina M. Gomez obtained her PhD with distinction with a thesis on the Redes de Trueque, the Community and Complementary Currency Systems in Argentina. She has published widely on complementary currencies, monetary innovation, solidarity finance and local economic development. She supervises doctoral candidates on monetary innovation, local economic development in Latin America, social and solidarity finance, grassroots economics and institutions. She is a member of the editorial boards of the journal *Development and Change* and the *International Journal of Community Currency Research*, and president of the Research Association on Monetary Innovation and Complementary Currency Systems. She launched a postgraduate diploma on Sustainable Local

Economic Development, a Massive Open Online Course on Local Economic Development and the Master in Local Development Studies at the Universidad de los Llanos in Villavicencio, Colombia. She has led development projects in Brazil, Nicaragua, Western Africa and Colombia funded by The World Bank, Woord en Daad and other cooperation agencies.

Pedro Goulart holds a PhD from Erasmus University Rotterdam. At Universidade de Lisboa, he is Deputy Director at CAPP, "Excellent" research centre with over 60 doctors, and assistant professor with tenure at ISCSP. Pedro is a specialist on child labour and education, economics for non-economists and local development, themes on which he led several funded research projects. His over 40 publications include articles in *Economics of Education Review*, *Journal of Educational Research*, *European Urban & Regional Studies* or *Social Science History*. He has been consulted by media, government and academic peers (e.g. CPLP, PETI, ZEW). He is also Portugal's EADI representative.

Alan Harding is Chief Economic Adviser to the Greater Manchester Combined Authority and a visiting professor at the Institute of Innovation Research at the Alliance Manchester Business School. He oversees the analytical work that supports Greater Manchester research, strategy development and evaluation. Previously, he spent 30 years in academia leading specialist research institutes and centres focusing on urban-regional development, policy and governance.

Rolph van der Hoeven is Professor (Emeritus since 2015) of Employment and Development Economics at the International Institute of Social Studies (ISS) of Erasmus University (EUR) in the Hague and member of the Committee for Development Planning of the United Nations. He has worked from 1974 till 2008 in different UN functions amongst others as Chief Economist at UNICEF, as Director for Policy Coherence at the ILO and as Scientific Manager at the World Commission on the Social Dimension of Globalization.

Daniel Horgos studied quantitative economics at the University of Tübingen and holds a PhD from Helmut Schmidt University. He received a Marie-Curie Scholarship and worked as post-doc at the Centro Studi Luca d'Agliano in Milan. His research interests focus on international economics, macroeconomics and subjective well-being related to globalization. As Director of Research he is currently heading the Research Unit at

MPC Capital and its investment vehicle MPC Container Ships. His research activities focus on macroeconomic market analyses for global maritime trade.

Edna Kabala is a dynamic and seasons applied economist working as a lecturer at The Copperbelt University. She holds BA in Economics and Development Studies from the University of Zambia and MA in Economics from the University of Botswana. Edna is currently in her final year of PhD studies in Economics at the University of Zambia where she is conducting a doctoral research titled "Mobile Money, Financial Inclusion and Livelihoods of Agents in Zambia." Edna is enthusiastic about research in international economics, economic development, financial inclusion and poverty reduction.

D'Artis Kancs is a Senior Economist and Lead Scientist at the European Commission, Directorate-General Joint Research Centre. He has been leading several research teams at the European Commission, including the macroeconomic modelling group in Sevilla, Spain, and the European Commission's Centre on Modelling in Ispra, Italy. Before joining the European Commission, he has held various teaching and research positions in universities and research institutes in the UK, Germany, Austria, USA and Belgium. He received PhD from the London School of Economics.

Peter Knorringa is Professor of Private Sector & Development at the International Institute of Social Studies (ISS) of Erasmus University Rotterdam. His research focuses on how entrepreneurs and firms in developing countries impact upon attempts to achieve more sustainable forms of development. He has worked on small- and medium-sized firms, entrepreneurship, local economic development, industrial clusters, role of trust and networking, global value chains, private governance and sustainability standards and more recently on frugal innovations. He is one of the co-founders and the present Academic Director of the Leiden-Delft-Erasmus Centre on Frugal Innovation in Africa, based in the Netherlands.

Luis Felipe López-Calva is the UNDP Regional Director for Latin America and the Caribbean since September 2018. He has nearly 30 years of professional experience, advising several Mexican governments, in addition to UNDP and most recently the World Bank where he most recently served as Practice Manager of the Poverty and Equity Global Practice (Europe and Central Asia). He was the co-director and lead author of the

World Development Report 2017 on "Governance and the Law." He was previously Lead Economist and Regional Poverty Advisor in the Bank's Europe and Central Asia Region, and Lead Economist at the Poverty, Equity and Gender Unit in the Latin America and Caribbean PREM Directorate, also at the World Bank. From 2007–2010, he served as Chief Economist for Latin America and the Caribbean at UNDP-RBLAC in New York. López-Calva has been associate editor of the *Journal of Human Development and Capabilities* and he is Fellow of the Human Development and Capabilities Association. He has also been Chair of the Network on Inequality and Poverty in the Latin America and Caribbean Economic Association. His research interests focus on labour markets, poverty and inequality, institutions and microeconomics of development, having presented his research at top institutions including Harvard University, Stanford University, University of California, Berkeley, University of California, San Diego, and the Organization for Economic Cooperation and Development Centre. He holds a master's degree in economics from Boston University, as well as a master's and a doctorate in economics from Cornell University.

Iga Magda is associate professor at the Warsaw School of Economics and vice president of the Institute for Structural Research (IBS) in Warsaw. She is also IZA Research Fellow. Her work is centred on labour economics, in particular wage and income inequalities, gender gaps, family policy and trade unions. She has published in *Journal of Comparative Economics*, *International Labour and Relations Review* and *Feminist Economics*, among others.

Miguel Martinez Lucio is Professor of International HRM and Comparative Industrial Relations at the University of Manchester where he is co-director of the Work & Equalities Institute (WEI). He is on the editorial board of the *European Journal of Industrial Relations* and *Critical Perspectives in International Business*. He is also joint chief editor of *New Technology, Work and Employment*. He has been an advisor and evaluator for various trade unions including UNISON, UNITE, CWU, UCATT the TUC and ETUC. He is a founder member of Critical Labour Studies and a member of the trade unions UNITE and UCU.

Anna Maria Mayda is Professor of Economics at Georgetown University, with a joint appointment in the Economics Department and School of Foreign Service. She studied statistics and economics at University of

Rome La Sapienza, where she received her degree summa cum laude. Before graduate school, she worked at the World Bank in the Latin America and Caribbean Region Unit. She received an MA and PhD in Economics from Harvard University, where she was also a doctoral fellow at the Center for International Development. She was a visiting scholar in several institutions including the Trade Unit of the IMF Research Department, University of Milan, EIEF in Rome and CEPII in Paris. More recently she was Senior Economist and Senior Adviser in the Office of the Chief Economist at the US State Department in the Obama Administration. She is a Research Affiliate at CEPR and IZA. Anna Maria Mayda's research mainly focuses on issues of trade, immigration and development economics and has been published in journals such as the *Quarterly Journal of Economics*, the *Review of Economics and Statistics*, the *Journal of International Economics*, *American Economic Journal: Applied and the European Economic Review*. She has also been awarded two National Science Foundation (NSF) grants. In terms of topics, she has worked on the determinants of individual attitudes towards trade and immigration across countries; the role played by interest groups in shaping US trade and migration policy; multilateral trade negotiations and preferential trade agreements; the determinants of international migration flows; the H-1B visa and the Refugee Resettlement programmes in the United States; the political and fiscal impacts of immigration to the United States.

Eduardo Ortiz-Juarez is Lecturer of Development Economics in the Department of International Development of King's College London, with a particular interest in the study of poverty dynamics, inequalities, social policy and green development. He is also a Senior Economic Research Consultant in the Strategic Policy Unit at the United Nations Development Programme (UNDP) and holds a research associate position at the Commitment to Equity Institute (CEQI), Tulane University. In 2020, he was the recipient of Mexico's National Prize of Public Finance awarded by the Chamber of Deputies' Centre for Public Finance Studies. Eduardo has 15 years of experience in international research and policy. He has served as Economist and Senior Economist in the Latin America and the Caribbean Bureau at the UNDP in New York, as Deputy Director of Economic and Social Analysis at the Mexican Ministry of Social Development, and has done extensive consultancy work for international organizations and national governments, including the World Bank, the Inter-American Development Bank (IDB), the United Nations University

World Institute for Development Economics Research (UNU-WIDER) and the Ibero-American General Secretariat (SEGIB).

Iva Peša is Assistant Professor of Contemporary History at the University of Groningen in the Netherlands. Her research is focused on the social and environmental history of Central Africa. Based on long-term fieldwork in Zambia and the Democratic Republic of Congo, she has published most recently on the environmental history of copper mining. Her work has appeared in journals such as *Environment and History* and the *Extractive Industries and Society*.

André Portela is Full Professor of Public Policy at the São Paulo School of Economics at Getulio Vargas Foundation (FGV EESP). He is the coordinator of the Center for Applied Microeconomics Studies and the Director of the Regional Center for Learning on Evaluation and Results for Brazil and Lusophone Africa at FGV EESP (FGV EESP C-Micro Clear). He is a researcher at the National Council of Research (CNPq) of Brazil. He holds a PhD in Economics from Cornell University, a Master's Degree in Economics from the University of São Paulo and a Bachelor's Degree in Economics from the Federal University of Bahia. He has been an Adjoint Associate Professor of Economics at Vanderbilt University and a visiting assistant professor at Cornell University. He develops research on Brazilian labour markets, education and evaluations of public policy.

Chiara Pronzato is associate professor at the University of Turin (Italy) and Director of the Impact Evaluation Unit at Collegio Carlo Alberto. [She studied Statistics at the University of Turin and then moved to ISER, University of Essex, where she obtained a PhD in Economics. In 2008, she was back to Italy with a post-doc position at Dondena, Bocconi University.] Her main research interests are in the fields of economics of the family, demography and policy evaluation methods. Her research appears in the *European Journal of Population*, *Journal of Population Economics*, *Review of Economics of the Household*, *Regional Studies*, *Economic Journal*, *Economic European Review* and *Management Science*.

Roberta Rabellotti is Professor of Economics at the University of Pavia. She regularly provides advice to the European Commission, the Inter-American Investment Bank (IADB), ILO, OECD; UNIDO; ECLAC-UN, UNCTAD. Her research focuses on innovation in developing countries, clusters, Global Value Chains and multinationals. She has published widely

in international journals and her last books were published by Harvard University Press, Edward Elgar, Palgrave and Routledge.

Anthony Rafferty is Professor of Employment Studies in the People, Management and Organisations (PMO) Division of AMBS and Managing Director of the Work and Equalities Institute (WEI) at the University of Manchester. His current research chiefly concerns issues surrounding labour market inequalities (e.g. gender and racial discrimination) and understanding the implications for organizational and management practice of issues of sustainable and inclusive growth. He is a member of the Economic and Social Research Council (ESRC) Peer Review College and has undertaken advisory roles or research consultancy for a variety of organizations including the UN International Labour Organisation (ILO), the European Institute for Gender Equality (EIGE), the UK Commission for Skills and Employment (UKCES), The European Foundation (Eurofound) and the UK Research Council. He has acted as an invited expert advisor to the UK government on employment and welfare reform issues giving oral evidence to the House of Commons Work and Pensions Select Committee.

Raul Ramos is a professor at the Department of Econometrics, Statistics and Applied Economics of the University of Barcelona, researcher at AQR-IREA, IZA Research Fellow and GLO Fellow. He acts as Vice-Dean of Research and Doctorate of the Faculty of Economics and Business of the University of Barcelona. His research interests include labour market analysis from both regional and international perspectives, with a specific focus on the impacts of globalization, European integration and migration on wage inequality and unemployment. He is coeditor of the *European Journal of Development Research*, Director of *Revista de Economía Laboral* and associate editor of *Regional Studies, Regional Science*.

James Rees is reader and Deputy Director of ICRD. He joined the University of Wolverhampton in Summer 2019; prior to that he was at the Open University and also worked at the Universities of Birmingham and Manchester in a range of disciplines including Geography, Politics and Social Policy. His work is notably inter-disciplinary and is concerned with the voluntary sector in its broadest sense, but also more specifically in its relationship with current transformations in public services in the UK. At the Third Sector Research Centre, University of Birmingham, he led the public service delivery programme, research from which helped set the

agenda for debates on the involvement of voluntary organizations in public services.

Nita Rudra is Professor of Government at Georgetown University. Her research interests include the politics of globalization, trade, foreign investment, development, democracy, inequality, taxation and redistribution. Her work appears in the *British Journal of Political Science*, *Journal of Politics*, *American Journal of Political Science* and *Comparative Political Studies, International Organization and International Studies Quarterly.* Her most recent book (with Ida Bastiaens) is titled: *Democracies in Peril.* She has been a recipient of the Woodrow Wilson International Center for Scholars Fellowship, the Fulbright-Nehru Foundation Academic Fellowship and the International Affairs Fellowship by the Council on Foreign Relations.

Katarzyna Sałach is a PhD candidate in Economics at the University of Warsaw, Poland. Her research interests include income and wealth inequality, wage differentials, poverty and energy poverty, labour economics and political economy. She has served as a consultant to the World Bank. Born in Poland, she has obtained her BSc in Mathematics and MA in Econometrics and Computer Science from the University of Warsaw.

AnnaLee Saxenian is a professor in the School of Information at the University of California, Berkeley. Her scholarship focuses on regional economies and the conditions under which people, ideas and geographies combine and connect into hubs of economic activity. She was Dean of the School of Information from 2004 to 2019, and upon stepping down she received the Berkeley Citation "for distinguished achievement and notable service to the University." She has served as a member of the Apple Academic Advisory Board, and Chair of the Advisory Committee for the National Science Foundation Division of Social, Behavioral, and Economic Sciences (NSF-SBE). She has published widely in journals of economic geography, regional economic development and industrial change. She holds a PhD from MIT, a master's from U.C. Berkeley and a BA from Williams College.

Marianne Sensier is an Economist and Research Fellow at the Alliance Manchester Business School. Marianne studied at the University of Sheffield and has worked at the University of Oxford and in Economics at the University of Manchester. Marianne has worked on a number of ESRC, Leverhulme Trust and ESPON funded research projects. Marianne

has written blogs for Policy@Manchester and her research areas include regional business cycles, inequalities and economic resilience**.**

Boriss Siliverstovs is a Principal Research Economist at the Bank of Latvia and a research fellow at the KOF Swiss Economic Institute at ETH Zurich. He also worked as a Senior Economist at the KOF Swiss Economic Institute at ETH Zurich and DIW Berlin. He holds his PhD degree from the University of Aarhus, Denmark. His main interests are nowcasting, real-time datasets and data analysis. He published more than 60 articles in peer-refereed journals in such diverse areas as forecasting, mixed-frequency data, migration, housing market, natural gas market, gravity models of trade and climate change.

James P. Smith joined Rose Li and Associates in 2020. He was Senior Economist at RAND from 1977 to 2020. He has been principal investigator of numerous National Institutes of Health funded projects, including studies of immigration, economics of aging, wealth accumulation, relation of health and economic status and impacts of early life on later life outcomes. He serves as key advisor to all HRS sister surveys. He was elected to National Academy of Medicine in 2011. He holds a PhD in Economics from University of Chicago, and honorary doctorates from University of Sterling and Trinity College Dublin.

Oded Stark is Distinguished Fellow at the Center for Development Research, University of Bonn, and Distinguished Professor at the University of Warsaw. He served as Adjunct Professor at the University of Tuebingen, Distinguished Research Scholar at Georgetown University, University Professor (Chair in Economic and Regional Policy) at the University of Klagenfurt, Honorary University Professor of Economics at the University of Vienna, Professor of Economics (Chair in Development Economics) at the University of Oslo, and as Professor of Population and Economics and as the Director of the Migration and Development Program at Harvard University. He has written on applied microeconomic theory, development economics, population economics, the economics of migration, labor economics, evolutionary economics, urban economics, regional economics, welfare economics, risk-taking behavior, and the theory of the firm. He is the author of the critically acclaimed books *The Migration of Labor* (Oxford and Cambridge, MA: Blackwell 1991 and 1993), and *Altruism and Beyond, An Economic Analysis of Transfers and Exchanges Within Families and Groups* (Cambridge: Cambridge University

Press 1995 and 1999), and is the co-editor of the *Handbook of Population and Family Economics* (in *Handbooks in Economics*; Amsterdam: North-Holland 1997, and Beijing: Economic Science Press 2004). Oded Stark is Doctor honoris causa (University of Warsaw), a Humboldt Awardee, a Ministry of Science and Higher Education (Poland) Lifetime Achievement Awardee, and a Presidential Professor of Economics (Poland).

Finn Tarp is Professor of Economics, University of Copenhagen, and coordinator of the University of Copenhagen Development Economics Research Group (DERG). The Director of UNU-WIDER from 2009–18, he is a leading international expert on development strategy and foreign aid. He has published widely in top economics and specialized field journals and is the editor of a wide range of book volumes with leading academic publishers. In addition to his university posts, he has held senior appointments and advisory positions with international organizations and several national governments and is a member of a series of international committees and advisory bodies.

Lara Patrício Tavares is an assistant professor at Instituto Superior de Ciências Sociais e Políticas, Universidade de Lisboa, where she also acts as deputy director of Centro de Administração e Políticas Públicas. Lara is a social scientist with a formal background in Economics—bachelor and master from Universidade Nova de Lisboa, and PhD from Institute for Social and Economic Research, University of Essex (UK). Her research interests include population economics and social demography. Lara's work was published in the journals *Population and Development Review, European Journal of Population, Social Science Research, Population Review, Health Policy, Journal of International Migration and Integration* and *Revista de Historia Transportes, Servicios y Telecomunicaciones.*

Graciela Teruel has been a researcher in the Department of Economics at Universidad Iberoamericana since 1998. In 2016 she was appointed director of the Research Institute for Equitable Development (EQUIDE). Her research includes impact evaluation of social programmes, including large-scale ones, such as Seguro Popular and Prospera; migration; changes in the well-being of individuals and the measurement of poverty. She is the principal investigator, with Luis Rubalcava, of the Mexican Family Life Survey and member of the National Council for the Evaluation of Social Policy in Mexico from 2006–2020. She holds a PhD in Economics from the University of California, Los Angeles.

Duncan Thomas is the Norb F. Schaefer Distinguished Professor of International Studies and Professor of Economics at Duke University. He contributed to the design, collection and analysis of the Indonesia Family Life Survey along with other longitudinal studies. His research investigates the inter-relationships between health, human capital and socio-economic status over the life course, highlighting dynamics within households and families and the role of the environment. He holds a PhD from Princeton University. He is a research associate at the National Bureau of Economic Research and Fellow of the Bureau for Research in the Economic Analysis of Development.

Ce Wu received her PhD in Agricultural Economics with Best Dissertation Award from Purdue University. Her research is specialized on economic development and international trade issues in Asia and Africa. She has taught economics and MBA courses at College of DuPage and Indiana Wesleyan University.

List of Figures

The Post-Cold War Era and the Labour Market: An Overview

Labour Without Borders

International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows

International Competition Intensified: Job Satisfaction Sacrificed?

Gender Pay Gaps in Domestic and Foreign-Owned Firms

Deglobalization and Labour: A New Era?

Technological Change and the Future of Work

Are Workers in the Developing World Winners or Losers in the Current Era of Globalization?

The Role of Labour in Capability Upgrading: The Case of Emerging Market Multinationals

Employment Effect of Innovation

Structural Transformation, Biased Technological Change and Employment in Vietnam

Is the Canary Alive and Kicking? Labour's Voice Under Crises

Labour and the State in India: Casualisation as Reform

Wages, Employment, and Economic Shocks: Evidence from Indonesia

The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach

A Vulnerability Approach to the Definition of the Middle Class

Disaster Disparities and differential Recovery in New Orleans

List of Tables

International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows

International Competition Intensified: Job Satisfaction Sacrificed?

Gender Pay Gaps in Domestic and Foreign-Owned Firms

Are Workers in the Developing World Winners or Losers in the Current Era of Globalization?

Employment Effect of Innovation

Structural Transformation, Biased Technological Change and Employment in Vietnam

'Helping a Large Number of People Become a Little Less Poor': The Logic of Survival Entrepreneurs

The Developmental Potential of Frugal Innovation Among Mobile Money Agents in Kitwe, Zambia

Is the Canary Alive and Kicking? Labour's Voice Under Crises

Wages, Employment, and Economic Shocks: Evidence from Indonesia

Growth and Recession: Underemployment and the Labour Market in the North of England

The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach

A Vulnerability Approach to the Definition of the Middle Class

Disaster Disparities and Differential Recovery in New Orleans

The Post-Cold War Era and the Labour Market: An Overview

Pedro Goulart, Raul Ramos, and Gianluca Ferrittu

Work, and not idleness, is the indispensable condition of happiness for every human being
—Leo Tolstoy (1892), *A Dialogue Among Clever People*

Nobody who works hard should be poor
—David K. Shipler (2005), *The Working Poor: Invisible in America, Random House*

P. Goulart (✉)
CAPP, Institute of Social and Political Sciences, Universidade de Lisboa, Lisbon, Portugal
e-mail: pgoulart@edu.ulisboa.pt

R. Ramos
Department of Econometrics, Statistics and Applied Economics, AQR-IREA, University of Barcelona, Barcelona, Spain
e-mail: rramos@ub.edu

G. Ferrittu
Lisbon School of Economics and Management - ISEG, Universidade de Lisboa, Lisbon, Portugal
e-mail: gferrittu@phd.iseg.ulisboa.pt

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_1

We have become rich countries of poor people
—Joseph Stiglitz (2006), *The Financial Times*

Imagine the uproar when Uber's cars start arriving without drivers.
—Martin Ford (2015), *Rise of the Robots: Technology and the Threat of a Jobless Future, Perseus Books*

1 A New Economic Era: From the Post-War Economic Boom to Secular Stagnation?

The economic growth performance in the post-Second World War era emerged as spectacular compared to previous periods. Using historical data from Bolt & van Zanden (2020), 43 years would have been required to double world per capita income with the average annual growth rate of 1.7% observed between 1900 and 1949. However, between 1950 and 1989, real per capita income grew at an average value of 2.6% per year, implying a doubling of income every 28 years. A similar trend has been observed during the last three decades, even taking into account that the Great Recession caused a sharp decrease in output in most economies. Real GDP per capita has multiplied by a factor of five between 1950 and 2019. As highlighted by Chau & Kanbur (2018), "before this period it took a thousand years for world per capita GDP to multiply by a factor of fifteen."

The post-war economic boom (also known as Golden Age of Capitalism—Marglin & Schor, 1992) was characterised by twenty years of sustained economic growth with high levels of labour productivity and low unemployment levels in West capitalist countries under a Keynesian policy framework. However, in the 1970s when the Bretton Woods monetary system collapsed and the sharp rise in oil prices in 1973, the reaction of Western Central Banks cutting interest rates to encourage growth ended up in an unexpected stagflation with high levels of inflation and unemployment at the same time (Bruno & Sachs, 1985). The fight against inflation under the new monetarist paradigm (Romer & Romer, 2013) and the structural changes after the oil crisis brought the world economy back to economic growth. In fact, the 1990s started with a mild recession followed by a robust period of expansion during which high productivity and income growth returned to most economies, a trend that was briefly interrupted at the beginning of the 2000s but followed by the most severe economic downturn since the 1929's Great Depression.

The Great Recession that occurred between 2007 and 2009 was not felt equally around the world; whereas most of the developed countries fell into a severe and long-lasting recession, other economies such as China or India grew substantially during this period (Keeley & Love, 2010). Moreover, the recovery after this crisis has not been as strong as expected. Summers (2018) revived the secular stagnation hypothesis to describe the economic situation in the United States after the Great Recession. The secular stagnation idea was formulated by Hansen (1939) to describe the macroeconomic developments of the US economy after the Great Depression. He identified a variety of factors that could explain the low investment demand in the 1930s, mainly an ageing population, too little immigration and reduced technological progress, factors that according to Summers could have also been relevant in the 2010s. Although the debate among economists regarding this issue is not fully closed, in the meanwhile the world economy has faced a new global shock associated with the COVID-19 pandemics.

2 A New Political Framework

While the period after the end of Second World War was characterised by decolonisation (also followed by an increase in the number of democracies), the beginning of the nineties was marked by the fall of the Berlin Wall and the official end to the Cold War, implying a major change in international and national institutional landscapes. In Europe, this meant a political approximation of countries with an expanding European Union and the expansion of liberal democracies. However, this period was also marked by a retreat of the state and a general intellectual move towards a more market-based economy. While the economy generally benefited, the position of labour became considerably frailer.

In developed countries, following the earlier elections of Thatcher and Reagan and contributing to the declining formal labour protection and falling union rates (Cummins, 2019). But also, development theorists and governments, often propelled by what would be coined as the "Washington Consensus," were walking towards development paths based in markets and less state intervention (Gore, 2000) with other authors such as Rodrik (2009) defending that economic development usually requires following policies that are tailored to local economic and political conditions rather than following "one common recipe." At the same time, the position of

labour showed, at best, contradictory signs or even became considerably frailer.

These major changes in international and national institutional landscapes influenced the world of labour also by declining formal labour protections in developed countries, putting national labour movements under severe pressure, and by widening the inequality in the Global South and neglecting labour markets in developing countries. A more globalised and interconnected world also made inequalities more visible, discomforting and conflictive. In this sense, labour institutions and agencies were challenged throughout all this period, and intersectionality often comes to make more severe the situation of some (Berg & Kucera, 2008).

3 Difficult Times for Labour Markets

The intersection between macroeconomic developments and policy paradigms has clearly determined the evolution of labour markets. From this perspective, it is important to start recognising that, although there are still significant differences among (and within) countries, the post-war period has seen a global major improvement in living and working conditions accompanied by several transformations—see, for instance, Papastefanaki & Potamianos (2020) for Southern Europe or UN-Habitat (2006) on structural changes in developing countries. Education has been expanded, increasing the qualification of the labour force. After a century of compulsory schooling in the leading countries, schooling has been democratised throughout the globe, with developing countries at different stages now educating most of their youths and reducing gender inequalities (Frankema, 2009), although there is still a long way to go (Patrinos & Psacharopoulos, 2011). With a global population of 7.7 billion inhabitants in 2020 according to the World Bank, the labour force has nearly reached 3.5 billion workers with participation rates of population aged over 15 around 60%, although showing a clear decreasing trend in the last decades (see Fig. 1). While the decrease is more intense for men than for women, a similar trend is observed across sexes. The main reason is the delayed transition of young adults from education to work. But the main challenges are related to population ageing and the decline of the labour force due to lower fertility rates, two trends that are clearly observed in developed countries since some decades ago (D'Addio et al., 2010) and that will affect developing countries quite soon (ILO, 2018). In this context, migration has been seen as a potential solution. In fact, the United

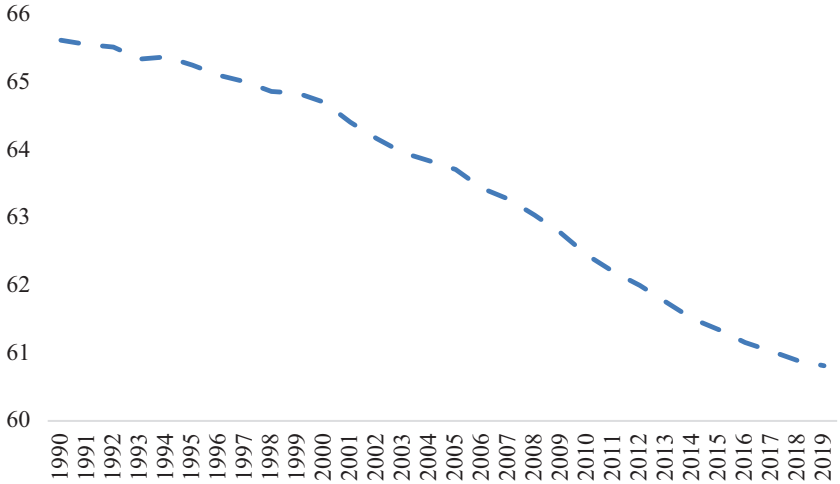

Fig. 1 Labour force participation rate (% of total population ages 15+). (Source: Own elaboration using data from the World Bank)

Nations (2000) report, where the migration volumes required to mitigate the effects of population decline and ageing were estimated, had a significant impact on the academic community (e.g. Card, 2005) and among citizens, but the debate about the pros and cons of this approach is still open.

As highlighted by Basso & Jimeno (2021) the second challenge that will shape the macroeconomy and the labour market in the coming decades is the technological change associated with the new wave of automation brought by developments in robotics and in artificial intelligence. In fact, if we look at the relative contribution of the factors behind real GDP per capita growth as shown in Fig. 2, we can see that capital and total factor productivity explain the highest part of global economic growth during the last thirty years. In fact, the last three decades have been marked by significant changes in the way global production is organised and robot adoptions can even accelerate the process (Cilekogl et al., 2021). Although the implementation of new technologies can be beneficial in the long run as it shifts the production frontier, it also involves a process of adaptation that has clear winners and losers. In fact, there is clear evidence about the relationship between technological change, outsourcing and job polarisation (Goos et al., 2014). Last, it is important to mention the potential

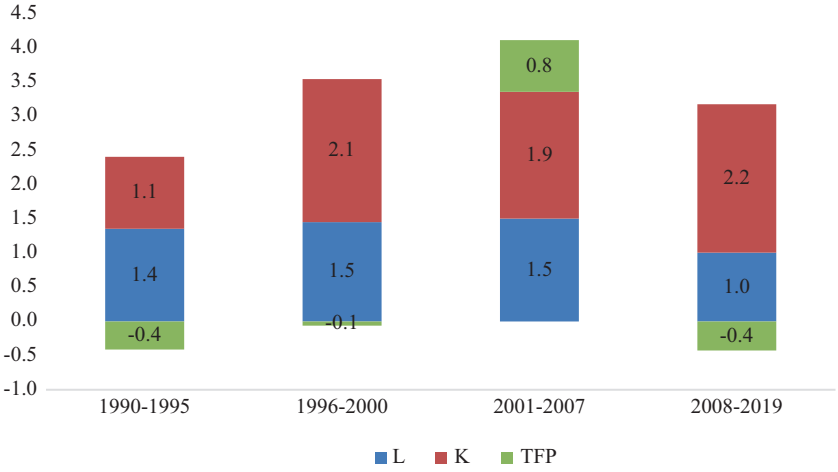

Fig. 2 Growth accounting of world real per capita GDP. (Source: Own elaboration using data from The Conference Board. Note: World real per capita GDP growth: 1990–1995: 2.0%, 1996–2000: 3.5%, 2001–2007: 4.1%, 2008–2019: 2.8%)

effects that the transition to a circular economy could have on labour markets (Méda, 2019). As Laubinger et al. (2020) highlight, structural changes in the economy from material-intensive to more labour-intensive activities could have potential benefits on employment, although the size of this effect is still unclear and difficult to quantify.

But the last decades have also seen important transformations in industrial relations and several labour market institutions have been reformed, particularly after the Great Recession (Adascalitei & Pignatti Morano, 2016). In the post-Cold War world, with declining unions and with the increase of what was once called atypical forms of work, agency is under challenge. While examining the US labour market outcomes, Groshen & Holzer (2021) describe the last 40-year trends as "disappointing," particularly for the working class. Would labour market policies be able to cope with the challenges ahead?

4 Concluding Remarks

During the last two years, the COVID-19 pandemic has accelerated some trends and challenges that were already affecting labour markets both in developed and developing countries. These include accelerated digitalisation and automation (partly related to the expansion of telework during lockdowns) and a growing relevance of non-standard forms of employment (such as platform workers) (ILO, 2021), but also migration responses to climatic change (Mcauliffe & Triandafillidou, 2021). The disruption caused by COVID-19 has negatively affected vulnerable and disadvantaged low-income households, migrants and workers in the informal sector, particularly in those countries with less-developed social safety nets making more evident the need to build stronger and more resilient institutions. But the pandemic has also led to a rise in protectionism and restrictions around the world that has clearly affected trade, foreign direct investment and migration (UNCTAD, 2020). However, as the economy recovers, the pandemic could also act as a catalyst for deep transformations in global value chains thanks to the enormous rescue and stimulus packages that have been adopted by developed economies. For instance, the twin transition (green and digital) that is the key element in the growth strategy of the European Union will undoubtedly change the economic landscape and the functioning of labour markets for the next decade. The contributions in the two volumes of this book can be helpful to achieve a better understanding of the present situation, but also the future developments.

References

Adascalitei, D., & Pignatti Morano, C. (2016). Drivers and Effects of Labour Market Reforms: Evidence from a Novel Policy Compendium. *IZA Journal of Labor Policy, 5*(15), 1–32.

Basso, H. S., & Jimeno, J. F. (2021). From Secular Stagnation to Robocalypse? Implications of Demographic and Technological Changes. *Journal of Monetary Economics, 117*, 833–847.

Berg, J., & Kucera, D. (2008). Labour Institutions in the Developing World: Historical and Theoretical Perspectives. In J. Berg & D. Kucera (Eds.), *Defence of Labour Market Institutions.* The International Labour Organization. Palgrave Macmillan.

Bolt, J., & van Zanden, J. L. (2020). *Maddison Style Estimates of the Evolution of the World Economy: A new 2020 update*. Maddison-Project Working Paper WP-15.

Bruno, M., & Sachs, J. D. (1985). *Economics of Worldwide Stagflation*. Harvard University Press.

Card, D. (2005). Is the New Immigration Really so Bad? *The Economic Journal, 115*(507), F300–F323.

Chau, N. H., & Kanbur, R. (2018). The Past, Present and Future of Economic Development. In *Towards a New Enlightenment? A Transcendent Decade*. BBVA.

Cilekogl, A., Moreno, R., & Ramos, R. (2021). *The Impact of Robot Adoption on Global Sourcing*. IZA DP 14255.

Cummins, I. D. (2019). Neoliberalism. In S. Pollock, K. Parkinson, & I. Cummins (Eds.), *Social Work and Society: Political and Ideological Perspectives* (pp. 50–62). Policy Press.

D'Addio, A. C., Keese, M., & Whitehouse, E. (2010). Population Ageing and Labour Markets. *Oxford Review of Economic Policy, 26*(4), 613–635.

Frankema, E. (2009). The Expansion of Mass Education in Twentieth Century Latin America: A Global Comparative Perspective. *Revista de Historia Económica, 27*(3), 359–396.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review, 104*(8), 2509–2526.

Gore, C. (2000). The Rise and Fall of the Washington Consensus as a Paradigm for Developing Countries. *World Development, 28*(5), 789–804.

Groshen, E. L., & Holzer, H. J. (2021). *Labor Market Trends and Outcomes: What Has Changed since the Great Recession?* IZA Policy Paper (173)

Hansen, A. H. (1939). Economic Progress and Declining Population Growth. *American Economic Review, 29*(1), 1–15.

ILO. (2018). *World Employment and Social Outlook. Trends 2018*. ILO.

ILO. (2021). *World Employment and Social Outlook 2021: The Role of Digital Labour Platforms in Transforming the World of Work*. ILO.

Keeley, B., & Love, P. (2010). *From Crisis to Recovery: The Causes, Course and Consequences of the Great Recession*. OECD Insights.

Laubinger, F., & Lanzi, E., & Chateau, J. (2020). *Labour Market Consequences of a Transition to a Circular Economy: A Review Paper*. OECD Environment Working Papers (162). OECD Publishing, Paris.

Marglin, S. A., & Schor, J. B. (Eds.). (1992). *The Golden Age of Capitalism: Reinterpreting the Postwar Experience*. Oxford University Press.

Mcauliffe, M., & Triandafillidou, A. (Eds.). (2021). *World Migration Report 2022*. International Organization for Migration (IOM).

Méda, D. (2019). Three Scenarios for the Future of Work. *International Labour Review, 158*, 627–652.

Papastefanaki, L., & Potamianos, N. (2020). *Labour History in the Semi-periphery. Southern Europe, 19th–20th Centuries*. De Gruyter Oldenbourg.

Patrinos, H. A., & Psacharopoulos, G. (2011). *Education: Past, Present and Future Global Challenges (English)*. Policy Research Working Paper WPS 5616. World Bank Group, Washington, DC.

Rodrik, D. (2009). *One Economics, Many Recipes. Globalization, Institutions, and Economic Growth*. Princeton University Press.

Romer, C. D., & Romer, D. H. (2013). The Most Dangerous Idea in Federal Reserve History: Monetary Policy Doesn't Matter. *American Economic Review, 103*(3), 55–60.

Summers, L. H. (2018). Secular Stagnation and Macroeconomic Policy. *IMF Economic Review, 66*, 226–250.

UNCTAD. (2020). *Impact of the COVID-19 Pandemic on Trade and Development: Transitioning to a New Normal*. United Nations Publications.

UN-Habitat. (2006). Structural Transformation in Developing Countries: Cross Regional Analysis. https://unhabitat.org/structural-transformation-in-developing-countries-cross-regional-analysis

United Nations. (2000). *Replacement Migration: Is it A Solution to Declining and Ageing Populations?* United Nations.

PART I

Globalisation and Migration

Labour Without Borders

Gianluca Ferrittu, Raul Ramos, and Pedro Goulart

1 Of Walls Fallen and Built: Globalisation and Labour Mobility

Migration has always been part of the history of human societies (see Lucassen & Lucassen, 1997). Yet it is only after the fall of the Berlin Wall, when the end of the Cold War seemed to have created new corridors for

G. Ferrittu (✉)
Lisbon School of Economics and Management - ISEG, Universidade de Lisboa, Lisbon, Portugal
e-mail: gferrittu@phd.iseg.ulisboa.pt

R. Ramos
Department of Econometrics, Statistics and Applied Economics, AQR-IREA, University of Barcelona, Barcelona, Spain
e-mail: rramos@ub.edu

P. Goulart
CAPP, Institute of Social and Political Sciences, Universidade de Lisboa, Lisbon, Portugal
e-mail: pgoulart@edu.ulisboa.pt

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_2

trasnational mobility, that international migration has become a central topic in the political debate (Castles, 2002). Since then, the consequential neoliberal globalisation has embedded processes as internationalisation of markets for goods and services, means of production, financial systems and technology, increasing capital mobility.[1] But as the absolute number of migrants was rising throughout the last three decades, new walls around the globe have been putting up to deter movements of people, while migration policies have become more "selective" (see de Haas et al., 2018). Migration seems thus to represent a unique paradox in the current international socioeconomic system, where capital, information and technology can move without restraints, while people, and so labour, cannot (Munck, 2008).

Since Adam Smith, economic thinkers have studied the free mobility and circulation of labour in both geographical as well as occupational terms. Traditionally, we envision migration as an individual act for income maximisation, caused by geographic differences in the supply of and demand for labour (see Lewis, 1954; Fei & Ranis, 1961; Todaro, 1969, as in Massey et al., 1993), where workers are in search of higher wages and better working conditions. However, this neoclassical view has been challenged over the years, and alternative and more comprehensive frameworks have been provided in the literature (see Stark & Bloom, 1985; Stark, 1991; Taylor, 1986; Piore, 1979; Hugo, 1981; Massey, 1990; Card, 2001; Ramos, 2016). Among labour and development economists, the complex link between migration and development has also been largely surveyed (see Lucas, 2005; de Haas, 2010), especially in terms of brain drain (Jolly & Seers, 1972; Stark, 2003); remittances and poverty alleviation (Adams & Page, 2005; Bayangos & Jansen, 2011; Sobiech, 2019); and personal aspirations, well-being and human development (Ramos, 2019; Dito et al., 2019; Rodan & Huijsmans, 2020). Further, the circulation of labour has also been studied in relation to foreign direct investment and trade, mainly in terms of complementarity and substitution (see, for instance, Federici and Giannetti, 2010; Ortega & Peri, 2014). However, the interdependences between globalisation, increased trade

[1] See also the definition of "Globalisation" reported in Eurostat, UN et al., 2002 (Glossary).

and international migration, and their impact on labour markets, remains relatively unexplored.

The 2030 Agenda[2] recognises the importance of labour migration (e.g. SDG 10.7; 8.8) and international trade (SDG targets 17.10, 17.11 and 17.12) for development through different targets. But there have been difficulties in separating between the social scientific and the political in understanding both (see Castles, 2002; Chang & Lee, 2011), and theoretical and policy advancements have thus been distressed. Whilst Massey et al. in 1998 already argued that the "post-industrial, post-Cold War world" needed a new comprehensive theory of migration (as reported in Castles, 2010), this seems still to be missing (see, for instance, de Haas, 2021).

2 International Trade Since 1991

By the end of the Cold War, the rise in trade has coincided with an increasingly interconnected international production system, and labour markets integration. International trade and foreign direct investment (FDI) have been the main defining features and key drivers of global value chains (see OECD, 2021). This integration of labour markets and globalisation of production has spread technology and innovations, and wage convergence across regions (see Giuliani et al., 2005; Pietrobelli & Rabellotti, 2011). But it also relocates production systems along complex and fragmented transnational value chains, impacting the characteristics and distribution of job tasks, and job quality and security. The globalisation of processes has also been related to decreasing wages of low-skilled labour, disruptive changes in the workforce composition (Becker et al., 2013) and rising wage inequality (Baumgarten, 2010).

Figure 1 depicts the evolution of global exports and net inflows of foreign direct investments as a share of world GDP for the period 1990–2019. The estimates show that between 1990 and 2007 (with the only exception of the downtrend reported between 2001 and 2004 for FDI contribution), trade and FDI followed similar increasing trends. However, after the Great Recession of 2008, diverging patterns are reported. The share of exports of goods and services starts to unsteadily decrease, from about 5 per cent in 2007 to a bit less than 2 per cent of global GDP in 2019.

[2] See the Sustainable Development Goals, UN (2021). Available at [https://www.globalgoals.org/]. Last access: 01/09/2021.

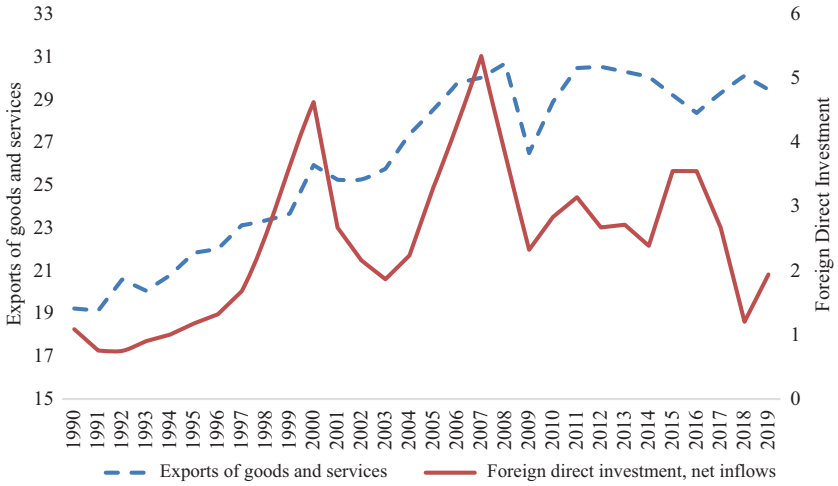

Fig. 1 World exports and FDI as a per cent of GDP. (Source: Authors' calculations based on World Bank data (World Bank, World Development Indicators, 2021b))

Although net inflows of FDI show a negative trend between 2007 and 2009, they also recover pre-crisis rates in a few years (in 2011 is already about 31 per cent of global GDP, as in 2007). Yet the estimates suggest trade is no longer growing after the Great Recession, while it is constant over the following reporting period. As outlined by Quiang and Kusek for the World Bank (2020), this may be related to economic drivers (e.g. declining rates of return on FDI), business factors, such as digital technology adoption and new organisational forms of international production, and policy factors, including investment uncertainties due to China and US trade war (see also Goulard, 2020), changes in US tax policy for capital repatriation and Brexit. Also increasing wages in emerging and new leading economies as China is likely to explain this trend (see Donaubauer & Dreger, 2016), which has also been considered a sign of "deglobalisation" (see van Bergeijk, 2019).

3 Labour Mobility in Distress

This globalised system of production has substituted the old pattern of labour, from the periphery to the core, with increasingly complex and diversified mobility of workforce between regions and countries (Saxenian, 2022). It also resulted in more diversified flows of migrants across borders through both legal and informal routes (see IOM, 2020). As such, nowadays "almost every country is an origin, host and transit country at the same time" (Kuptsch et al., 2019). However, as reported by Rapoport (2016), while globalisation has been firstly termed "the age of mass migration," the current international system seems globalised for everything except people. Despite the share of international migrants in the world population has slightly increased in the 1990s, as also did the absolute numbers of migrants according to UN estimates (increasing from 173 million in 2000 to 281 in 2020, see IOM's GMDAC[3]), the international migrants' stock has remained relatively stable over the last three decades, from 2.8 per cent in 1990 to almost 3.6 per cent in 2020 (op. cit.).

Yet the destinations of dominant migration are changing in relation to megatrends and economic development, and bilateral migration patterns require our attention. Figure 2 depicts the bilateral migration stocks as in 2017 for seven regional blocs, namely East Asia and Pacific, Europe and Central Asia, Latin America and The Caribbean, Middle East and North Africa, North America, South Asia, Sub-Saharan Africa.

According to the estimates reported, internal and interregional migration seem to be the most relevant types of migration, and South-South and North-North corridors are observed. The main migration within the same region is observed in Europe and Central Asia, with 61.8 million immigrants born in other countries of the same region (or 7 per cent of the total population in 2017). This is the case also for East Asia and Pacific (with 18.7 million immigrants from other East Asian and Pacific countries or about 1 per cent of the total population), Sub-Saharan Africa (20.4 million immigrants, 2 per cent of the population) and the Middle East and North Africa (13.7 million immigrants, or 3 per cent of the population). Relevant internal regional migrations are also detected in Latin America and the Caribbean, while from South Asia the main migration is majorly

[3] See the Migration data portal (2021). Available at [https://www.migrationdataportal.org/themes/international-migrant-stocks#key-trends-2000-2020]. Last access: 01/09/2021.

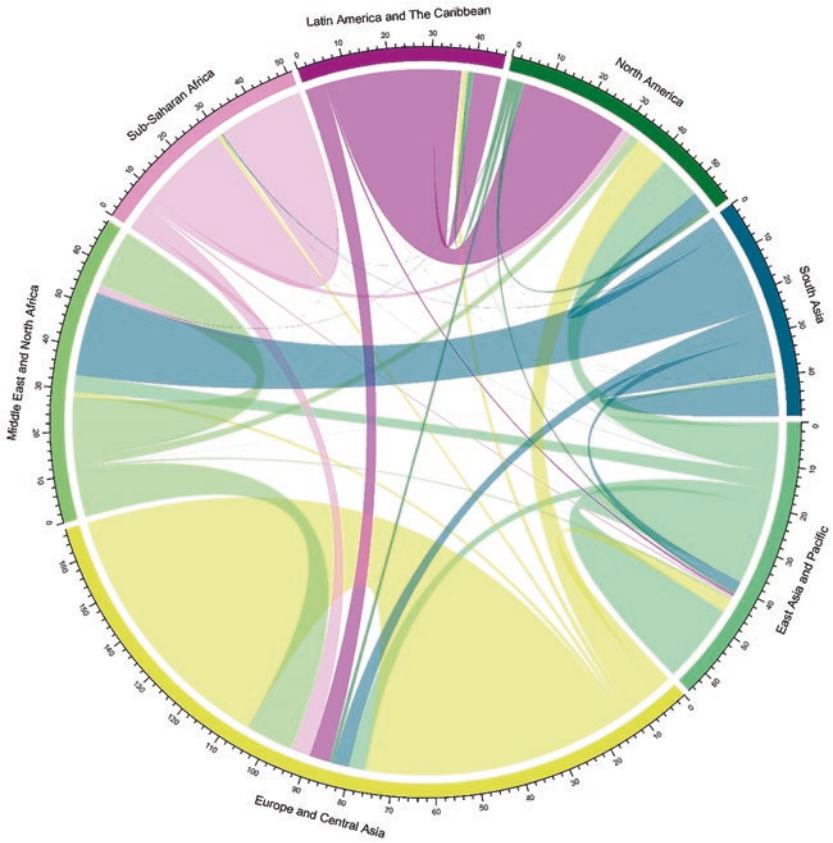

Fig. 2 Bilateral migrant stocks. (Source: Authors' calculations based on World Bank (2017). Note: The Chord diagram reports the number of people (in million) born in other countries in which they live at a given time of the year, by destination and by origin)

observed towards the Middle East and North Africa. This dynamic seems to reflect "the reality of globalisation as an uneven process with poles of development within the South promoting migration" (Munck, 2008). In terms of transregional migration, the figure also confirms the main corridor tends to be from developing countries to richer economies. These are the cases of Latin America and the Caribbean to North America (about

6.3 million immigrants registered), or the Middle East and North African and Sub-Saharan African migrants destined mainly to Europe and Central Asia (11.4 and 4.5 million immigrants, respectively). Further, the largest transregional corridor from East Asia and Pacific seems to be towards North America (with 10.3 million immigrants). According to UN estimates (see IOM's GMDAC, op. cit.), rural-to-urban migration has also been intensified over the last decades (since 2007, the global urban population has surpassed the rural population, op. cit.) contributing to circular and temporary migration (see also Dustmann & Görlach, 2016).

As people decide to migrate, they may also opt for remitting part of their income back to their countries of origin to support family and relatives (see IOM, 2006). Contributions of remittances to development at the macro and micro levels are thus well known in the literature (see Rapoport & Docquier, 2005; Bayangos, 2012). On the other hand, ambivalent impacts on economic growth and labour markets, such as increasing non-income labour and brain drain (Sutradhar, 2020), have also been surveyed.

Figure 3 reports the contribution of personal remittances received as a percentage of GDP per high- and low- and middle-income countries, and at the global level, over the last three decades. Since the 1990s, the estimates show remittances play a stable contribution for high-income countries. At the global level, it is possible to observe a general increasing trend of the contribution of remittances: whilst in the 1990 remittances were 0.4 per cent of GDP, in 2020 they have been increased to almost 1 per cent. This may be related to the steadily increasing role of remittances for middle-income countries. By the end of the Cold War, in this regional bloc, the share increased for almost 13 years, from about 1.1 per cent in 1990 to 1.8 per cent of GDP in 2003. After this year, the estimate shows a slight decrease until 2011 (1.4 per cent), to which it follows a period of general stagnation.

Regarding low-income countries, the contribution of remittances is higher in all the periods reported compared to the world average, and all the other regional blocs. While in the first decade a steady increase in the shares of remittances is detected, since 2003 we observe a sharp decline, from 4.2 in 2003 to 2.02 per cent of GDP in 2009. It follows a period of stagnation. The slump registered in remittances in 2003 allows for converging trends between low- and middle-income countries in the last years reported. In 2020, personal remittances in low- and middle-income countries were about 2.6 and 1.6 percent of GDP, respectively. In line with the

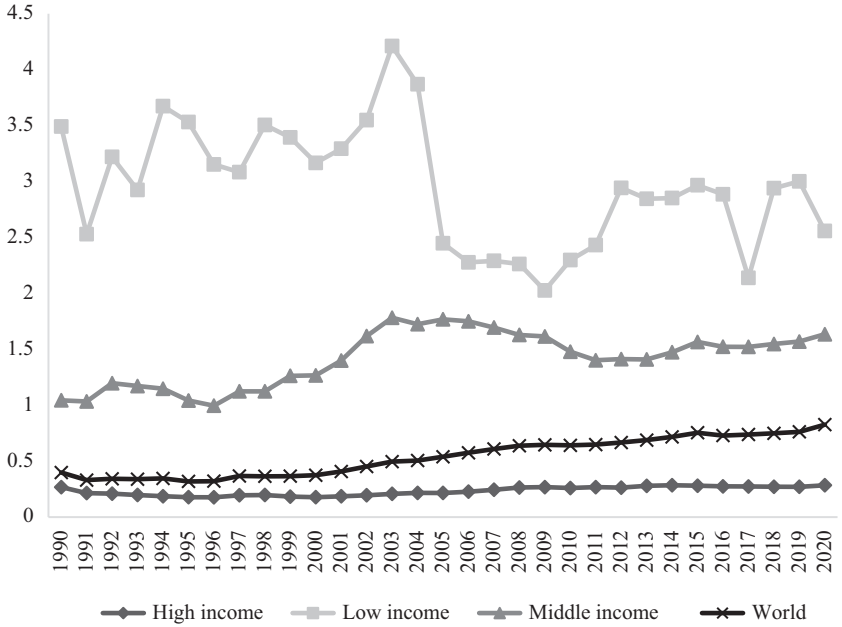

Fig. 3 Personal remittances, received (per cent of GDP). (Source: Authors' calculation based on World Bank data (World Bank, World Development Indicators, 2021a))

decay hypotheses (see Hunte, 2004), remittances stagnation detected after 2011 may be related to increases in household income in countries of origin, but also to cyclical recessions and economic shocks, stricter migration policies (see de Haas, 2005) and exchange rate flexibility (Lartey, 2016). Although the literature was dominated for a long time by the "brain drain" hypothesis (Beine et al., 2008), developing countries can also benefit from migration through other channels such as return migration (Wahba, 2021) and brain circulation (Saxenian, 2022). For instance, migrants acquire new skills that can be very valuable in their home countries, usually by setting up a business thanks to accumulated savings during their stay abroad (Mayr & Peri, 2009). Moreover, they can also transfer new ideas and norms that can contribute to economic development.

The last decades have also seen an interesting debate between economists regarding the impacts of immigration on host labour markets that

has also influenced the policy debate and citizen's attitudes towards migration. The earlier analysis by Card (1990) has been revisited by several renowned scholars such as Borjas (2003) and more recently, Borjas (2017) or Peri and Yasenov (2019). Although the debate is far from being closed, immigrants' positive effects on host labour markets seem to outweigh potential negative impacts.

The unfolding COVID-19 pandemic has temporarily slowed down the mobility of people, and the overall integration of production processes. Yet labour migrants still played a significant role in host countries in the response to the pandemic (see OECD, 2020a), while migration flows seem already to have re-started (see Bartolini, 2022). Remittances from developed to developing economies have diminished during the pandemic crisis, but they also proved to be resilient (World Bank, 2021).

4 In This Section: Globalisation and Migration

The first section of this volume, *Globalisation and Migration*, studies these issues by collecting relevant publications on the impacts of migration and globalisation of trade and labour markets on the world of work in the post-Cold War era.

In migration research, migrants decide to move according to economic and non-economic incentives. These involve supply-push factors (e.g. lack of employment and opportunities, natural disasters) and demand-pull factors, such as the host country's immigration policy, and more attractive job opportunities. Chapter 3 in this section, by Mayda, empirically investigates the determinants of bilateral immigration flows in 14 OECD countries by country of origin, by developing a model that puts greater emphasis on the demand-side of international migration. The author shows pull factors significantly increase the size of emigration rates, while the sign of the impact of push factors is rarely negative and, when it is, the size of the effect is not significant, and smaller than that of pull factors. Mayda (2022) suggests migration quotas are more binding than pull effects, and that migration quotas matter because they regulate the supply-side effects of migration flows. This is consistent with the evidence in the literature about the constraining role played by migration policies. While the author makes progress in explaining the determinants of international migration flows, this survey also leads to other interesting research questions of migration relative to other dimensions of globalisation, such as the possibility to apply common frameworks of empirical analysis for trade and migration.

In the literature, this has been complemented with novel comprehensive approaches on the determinants of flows of both capital and goods between countries (see for instance Beine et al., 2016).

Yet the line of demarcation between push and pull factors of migration may be blurry and not reflect the intrinsic complexity of individual migration decisions. As discussed by Stark (2022) in Chap. 4, received migration research has it that higher relative deprivation strengthens the incentive for people to migrate, and that often migration is a risky enterprise. Relative deprivation has been seen as a push factor in migration, and the level of risk involved in migration has been understood to reduce its attraction. Stark reveals a positive relationship between the level of relative deprivation experienced at origin and willingness to undertake risk-laden migration: he shows that higher relative deprivation is matched by riskier acceptable migration options. In expanding the range of acceptable risk-laden migration options, relative deprivation experienced at origin acts also as a migration pull factor.

The globalisation of labour mobility and production has largely transformed the world of labour in terms of skill levels (see Thoenig & Verdier, 2003). In line with ongoing demographic shifts and technological changes, future demand for labour migrants would have certain characteristics, in certain sectors and occupations (see OECD, 2020a, 2020b). These dynamics have substantial implications for the labour market, at both high- and low-skilled ends. With regard to the first, Chap. 5, by Faggian, studies the evolution of the main features and concepts of high-skilled migration, discussing potential future developments. While the categorisation of high-skilled migrants in the literature is not univocal, Faggian (2022) outlines that there is a strong justification for using education as a way to identify high-skilled migrants. The related determinants and consequences of this type of migration are then discussed in the chapter. The author argues big data may play a relevant role in shedding lights on high-skilled migrants, especially in terms of female migration. Yet this new availability of data also embeds complex challenges, such as too much emphasis on data-driven research, and data protection and privacy risks.

High-skilled migration has also been largely studied in terms of development implications. As previously mentioned, one of the main features of this literature is related to the phenomenon of brain drain, which focuses on the structural impoverishment of developing economies caused by the exodus of high-skilled labourers towards the richest countries. In contrast to this traditional approach, Chap. 6, by Saxenian, examines the role of international migration in promoting sectoral upgrading in migrants' origin countries. Particularly, the author surveys how Chinese and Indian-born engineers are accelerating the development of the information technology industries in their home countries. Saxenian suggests that this type of high-skilled labour mobility is likely to reverse the so-called brain drain, transforming it into "brain circulation," as individuals return home to establish business relationships, while maintaining their professional networks with advanced economies. Saxenian innovates by conceptualising migration of talents as a "reversible individual act." However, the author points out that this model, based on brain circulation, also raises questions about its validity outside the key regions of the Global South. Saxenian suggests that this dynamic seems to benefit only certain developing countries, which are characterised by high levels of political and economic stability, enough that immigrants will consider for returning home, and that have largely invested in higher education.

The increasing capital mobility related to markets integration promoted a rich debate about labour market implications (see, for instance, Acemoglu & Autor, 2011). But what about the effect of trade and offshoring on labourer's well-being and job satisfaction? Chapter 7 in this book, by Dluhosch and Horgos, considers the psychological consequences of globalisation of production in advanced economies. As the authors report in their analysis, off-shoring activities may generate a decline in the demand for low-skilled labour, as low-skill production tasks are imported from abroad. This may negatively impact low-skilled labourers by diminishing their income perspectives and promoting disruptive competition. Dluhosch and Horgos contribute to the literature by defining a model which captures "job satisfaction in conjunction with the income and distributional effects of offshoring." This is based on self-reported information about job satisfaction from the International Social Survey Program. Dluhosch and Horgos's (2022) results suggest some indicators of international trade confirm conventional emphases about disruptive effects of production internationalisation. However, they also show that job satisfaction is higher in countries with relatively high offshoring activities. This

would be possible because tasks change also in the home country's production, and employees are more satisfied with their work. The chapter thus sheds light on future research questions about the possible effects of other different non-pecuniary aspects of offshoring.

Various economic studies have also surveyed the effects of foreign direct investments and offshoring activities on workers' wages and employment (see Blomström et al., 1997; Poole, 2013) and gender inequality and pay gap (see Heyman et al., 2013). As outlined in Chap. 8 of this book, by Magda and Sałach (2022), in line with competition and trade theories, the gender pay gap should be smaller among foreign-owned companies than among domestically owned firms. This is because discrimination should be more costly for foreign-owned companies that are more likely to operate under competitive market conditions. At the same time, domestic firms are more likely to be domestic market-oriented, where gender-based pay discrimination is less costly. Yet the empirical evidence regarding these assumptions is inconclusive and still debated. In this section, Magda and Sałach contribute by analysing the differences in gender wage gaps between foreign and domestically owned firms in Poland. The authors find that the gender pay gap is larger in foreign-owned firms than in their domestic counterparts, but that the difference reported between the two ownership sectors is narrower than "raw data would suggest." The study also underlines that the foreign and the domestic sectors show different patterns of gender wage inequality. Magda and Sałach point out that, while the gender pay gap is larger within foreign-owned firms, wage premia mitigate the differences observed in this sector. At the same time, in domestic-owned firms, the gender pay gap within firm is lower compared to those that are foreign-owned, but firm policies seem to widen the female pay disadvantage at the sector level. The chapter thus sheds light on future research questions about potential concomitant drivers and interactions of the ownership sector's effects on gender pay gaps.

Within the process of internationalisation of labour markets, more vulnerability at the workplace and wider inequality is a risk. The evolution of migration policies is likely to play a major role in this transition. Kuptsch et al., for ILO, in 2019, were already asking "are we moving even further towards a global segmentation of labour markets along with persistent discriminatory practices against migrant workers?" (p. 7). It seems

countries would need to ask this question to design the future of labour mobility. Chapter 9 of this book, by Bartolini, contributes to this debate by summarising the major policy approaches to labour mobility expressed by the International Organization for Migration (IOM—UN Migration) over the past decades. While IOM has joined the United Nations system in 2016, becoming the Secretariat for the UN Migration Network, it has supported governments in managing migration since 1951.[4] In the chapter, the author systematically compiles the evolution of the organisational concepts and actions in relation to labour migration. Bartolini reviews to what extent IOM advocates and works to protect migrant workers, and to maximise the development potentials of labour mobility for both origin and destination countries, and the migrants themselves. According to the author, in line with IOM's labour mobility and human development agenda, novel predictable and equitable pathways to regular migration are still necessary for socioeconomic development and human security, especially in relation to the ongoing COVID-19 pandemic crisis.

Against the background of the currently unfolding pandemic crisis, and, before, the Great Recession, Chap. 10, by van Bergeijk and van der Hoeven, concludes this section by analysing the potential impact of deglobalisation on labour. According to the authors, after the Great Recession in 2008, international trade declined and protectionism increased, starting a process of deglobalisation of the international system of production (see also Fig. 1). This issue would have been reinforced by the current COVID-19 crisis. But while globalisation and deglobalisation are different phenomena, van Bergeijk and van der Hoeven (2022) argue the impact of the shocks on labour is similar. As such, some of the intrinsic problems of labour, such as informality and working poverty exacerbated the situation of the most vulnerable and the poorest, who remain the main victims of these processes. In the chapter, van Bergeijk and van der Hoeven suggest the context of the post-pandemic gives us a unique chance to design the future policies of global labour markets, which should take into account the current structural inequality between capital and labour, and the political and societal harm of segmentation in the labour market, without "retreating to protectionism."

[4] See IOM (2021). *IOM History*. Available at [https://www.iom.int/iom-history]. Last access: 01/09/2021.

References

Acemoglu, D., & Autor, D. H. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics, 4B*, 1043–1171.

Adams, R., & Page, J. (2005). Do International Migration and Remittance Reduce Poverty in Developing Countries? *World Development, 33*, 1645–1669.

Bartolini, L. (2022). An International Organization's Approach to Labour Migration for Human Development. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Baumgarten, D. (2010). Exporters and the Rise in Wage Inequality—Evidence from German Linked Employer-Employee Data. *Journal of International Economics, 90*, 201–217.

Bayangos, V. (2012). *Going with Remittances: The Case of the Philippines*. BSP Working Paper Series No. 2012-01. Quezon City: Bangko Central ng Pilipinas.

Bayangos, V., & Jansen, K. (2011). Remittances and Competitiveness: The Case of the Philippines. *World Development, 39*(10), 1834–1846. Elsevier.

Becker, S. O., Ekholm, K., & Muendler, M. A. (2013). Offshoring and the Onshore Composition of Tasks and Skills. *Journal of International Economics, 90*(1), 91–106.

Beine, M., Docquier, F., & Rapoport, H. (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *The Economic Journal, 118*(528), 631–652.

Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J. (2016). A Practitioners' Guide to Gravity Models of International Migration. *The World Economy, Special Issue: International Migration and Inequality Across Countries, 39*(4), 496–512.

Blomström, M., Fors, G., & Lipsey, R. E. (1997). Foreign Direct Investment and Employment: Home Country Experience in the United States and Sweden. *The Economic Journal, 107*(445), 1787–1797.

Borjas, G. J. (2003). The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics, 118*(4), 1335–1374.

Borjas, G. J. (2017). The Wage Impact of the Marielitos: A Reappraisal. *ILR Review, 70*(5), 1077.

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review, 43*(2), 245–257.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics, 19*(1), 22–64.

Castles, S. (2002). Migration and Community Formation under Conditions of Globalization. *International Migration Review, 36*(4), 1143–1168.

Castles, S. (2010). Understanding Global Migration: A Social Transformation Perspective. *Journal of Ethnic and Migration Studies, 36*, 1565–1586.

Chang, C. P., & Lee, C. C. (2011). Partisanship And International Trade: Some International Evidence. *Acta Oeconomica, 62*, 41–63.

de Haas, H. (2005). International Migration, Remittances and Development: Myths and Facts. *Third World Quarterly, 26*(8), 1269–1284.

de Haas, H. (2010). Migration and Development: A Theoretical Perspective 1. *The International Migration Review, 44*, 227–264.

de Haas, H. (2021). A Theory of Migration: The Aspirations-Capabilities Framework. *Comparative Migrations Studies, 9*(8), 1–35.

de Haas, H., Natter, K., & Vezzoli, S. (2018). Growing Restrictiveness or Changing Selection? The Nature and Evolution of Migration Policies. *International Migration Review, 52*(2), 324–367.

Dito, B., Mazzucato, V., White, A., & Veale, A. (2019). Transnational Migration, Health, and Well-Being: Nigerian Parents in Ireland and the Netherlands. *Comparative Migration Studies, 7*(44), 7–44.

Dluhosch, B., & Horgos, D. (2022). International Competition Intensified: Job Satisfaction Sacrificed? In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Donaubauer, J., & Dreger, C. (2016). *The End of Cheap Labour: Are Foreign Investors Leaving China?* DIW Berlin Discussion Paper No. 1598.

Dustmann, C., & Görlach, J. S. (2016). The Economics of Temporary Migrations. *Journal of Economic Literature, 54*(1), 98–136.

Eurostat, IMF, OECD, UN, UNCTAD, WTO. (2002). *Manual on Statistics of International Trade in Services, Annex II, Glossary.* United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations Conference on Trade and Development and World Trade Organization. New York: United Nations Publications.

Faggian, A. (2022). High-Skilled Migration: Past, Present, and Future. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Federici, D., & Giannetti, M. (2010). Temporary Migration and Foreign Direct Investment. *Open Economies Review, 21*, 293–308.

Fei, J. C. H., & Ranis, G. (1961). A Theory of Economic Development. *The American Economic Review, 51*(4), 533–565.

Foad, H. (2012). FDI and Immigration: A Regional Analysis. *The Annals of Regional Science, 49*, 237–259.

Giuliani, E., Pietrobelli, C., & Rabellotti, R. (2005). Upgrading in Global Value Chains: Lessons from Latin America Clusters. *World Development, 33*(4), 549–573.

Goulard, S. (2020). The Impact of the US–China Trade War on the European Union. *Global Journal of Emerging Market Economies, 12*, 56–68.

Heyman, F., Svaleryd, H., & Vlachos, J. (2013). Competition, Takeovers, and Gender Discrimination. *Industrial and Labor Relations Review, 66*(2), 409–432.

Huber, P., & Nowotny, K. (2020). Risk Aversion and the Willingness to Migrate in 30 Transition Countries. *Journal of Population Economics, 33*(4), 1463–1498.

Hugo, G. J. (1981). Village-Community Ties, Village Norms, and Ethnic and Social Networks: A Review of Evidence from the Third World. In D. J. F. Gordon & R. W. Gardner (Eds.), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*. Pergamon Press.

Hunte, C. K. (2004). Workers' Remittances, Remittance Decay and Financial Deepening in Developing Countries. *The American Economist, 48*(2), 82–94.

IOM. (2006). *Migrants' Remittances and Development*. International Organization for Migration—UN Migration.

IOM. (2020). *World Migration Report 2020*. International Organization for Migration—UN Migration.

Jolly, J., & Seers, D. (1972). The Brain Drain and the Development Process. In G. Ranis (Ed.), *The Gap Between Rich and Poor Nations* (chapter 15, pp. 365–379), International Economic Association Series. Palgrave Macmillan.

Kugler, M., & Rapoport, H. (2007). International Labor and Capital Flows: Complements or Substitutes? *Economics Letters, 94*, 155–162.

Kuptsch, U., Amara, D., Schmidt, L., Addati, P., & Walsh, K. (2019). *The Future of Work, The Centenary Initiative*. Issue Note Series No. 2, International Labour Office, Geneva.

Lartey, E. K. (2016). *Exchange Rate Flexibility and the Effect of Remittances on Economic Growth*. Policy Research Working Paper, No. 7932. World Bank: Washington, DC.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *Manchester School of Economic and Social Studies, 22*, 139–191.

Lucas, R. (2005). *International Migration and Economic Development: Lessons from Low-Income Countries*. Edward Elgar Pub.

Lucassen, J., & Lucassen, L. (1997). Introduction. In J. Lucassen & L. Lucassen (Eds.), *Migration, Migration History, History. Old Paradigms and New Perspectives*. Peter Lang.

Magda, I., & Salach, K. (2022). Gender Pay Gaps in Domestic and Foreign-Owned Firms. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Massey, D. S. (1990). Social Structure, Household Strategies, and the Cumulative Causation of Migration. *Population Index, 56*, 3–26.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Taylor, J. E., & Pellegrino, A. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review, 19*(3), 431–466.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1998). *Worlds in Motion: Understanding International Migration at the End of the Millennium*. Clarendon Press.

Mayda, A. M. (2022). International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Mayr, K., & Peri, P. (2009). Brain Drain and Brain Return: Theory and Application, to Eastern-Western Europe. *The B.E. Journal of Economic Analysis & Policy, 9*(1), Article 49.

Munck, R. (2008). Globalisation, Governance and Migration: An Introduction. *Third World Quarterly, 29*, 1227–1246.

OECD. (2020a). *Managing International Migration under COVID-19*. OECD Policy Responses to Covid-19. Organization for Economic Co-operation and Development (OECD), Paris.

OECD. (2020b). *COVID-19 and Key Workers: What Role Do Migrants Play in Your Region?* Organization for Economic Co-operation and Development (OECD).

OECD. (2021). *Enabling SME Linkages with Foreign Firms in Global Value Chains, in Middle East and North Africa Investment Policy Perspectives*. Organization for Economic Co-operation and Development (OECD).

Ortega, F., & Peri, G. (2014). Openness and Income: The Roles of Trade and Migration. *Journal of International Economics, 92*(2), 231–251.

Peri, G., & Yasenov, V. (2019). The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift. *Journal of Human Resources, 54*, 267–309.

Pietrobelli, C., & Rabellotti, R. (2011). Global Value Chains Meet Innovation Systems: Are There Learning Opportunities for Developing Countries? *World Development, 39*(7), 1261–1269.

Piore, M. J. (1979). *Birds of Passage: Migrant Labor in Industrial Societies*. Cambridge University Press.

Poole, J. (2013). Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility. *Review of Economics and Statistics, 95*(2), 393–406.

Ramos, R. (2016). *Gravity Models: A Tool for Migration Analysis*. IZA World of Labor, Institute for the Study of Labor (IZA), Bonn, Iss. 239

Ramos, R. (2019). Migration Aspirations among Youth in the Middle East and North Africa Region. *Journal of Geographical Systems, 21*(4), 487–507.

Rapoport, H. (2016). *Migration and Globalization: What's in It for Developing Countries?*, IZA Discussion Papers, No. 10109. Institute for the Study of Labor (IZA), Bonn.

Rapoport, H., & Docquier, F. (2005). *The Economics of Migrants' Remittances.* IZA Discussion Paper, No. 1531.

Rodan, L., & Huijsmans, R. B. C. (2020). 'Our Generation…' Aspiration, Desire, and Generation as Discourse Among Highly Educated, Portuguese, Post-austerity Migrants in London. Opens External. *European Journal of Development Research, 33*, 147–164.

Saxenian, A. (2022). From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Sobiech, I. (2019). Remittances, Finance and Growth: Does Financial Development Foster the Impact of Remittances on Economic Growth? *World Development, 113*, 44–59.

Stark, O. (1991). *The Migration of Labour*. Basil Blackwell.

Stark, O. (2003). *Rethinking the Brain Drain*. Discussion Papers, University of Bonn, Center for Development Research (ZEF).

Stark, O. (2022). A Unified Model of Relative Deprivation and Risk-Laden Migration. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Stark, O., & Bloom, D. E. (1985). The New Economics of Labor Migration. *American Economic Review, 75*(2), 173–178.

Sutradhar, S. R. (2020). The Impact of Remittances on Economic Growth in Bangladesh, India, Pakistan and Sri Lanka. *International Journal of Economic Policy Studies, 14*, 275–295.

Taylor, J. E. (1986). Differential Migration, Networks, Information and Risk. In O. Stark (Ed.), *Migration Theory, Human Capital and Development*. JAI Press.

Thoenig, M., & Verdier, T. (2003). A Theory of Defensive Skill-Biased Innovation and Globalization. *American Economic Review, 93*(3), 709–728.

Todaro, M. P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *American Economic Review, 59*, 138–148.

van Bergeijk, P. (2019). Deglobalization 2.0 Trade and Openness During the Great Depression and the Great Recession. In *Deglobalization 2.0 Trade and Openness During the Great Depression and the Great Recession*. Edward Elgar.

van Bergeijk, P., & van der Hoeven, R. (2022). Deglobalization and Labour: A New Era? In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Wahba, J. (2021). *Who Benefits from Return Migration to Developing Countries?* IZA World of Labor 2021: 123.

World Bank. (2020). *Global Investment Competitiveness Report 2019/2020: Rebuilding Investor Confidence in Times of Uncertainty*. World Bank Group.

World Bank. (2021). *Resilience COVID-19 Crisis through a Migration Lens*. Migration and Development Brief, N. 34. World Bank Group: Washington, DC.

World Bank, World Development Indicators. (2021a). *Personal Remittances, Received (% of GDP)*. Retrieved from https://data.worldbank.org/indicator/BX.TRF.PWKR.DT.GD.ZS

World Bank, World Development Indicators. (2021b). *World Exports and FDI as a % of GDP*. Retrieved from https://data.worldbank.org/indicator/BX.KLT.DINV.WD.GD.ZS and https://data.worldbank.org/indicator/NE.EXP.GNFS.ZS

International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows

Anna Maria Mayda

1 Introduction

International migration patterns vary considerably over time, and across destination and origin countries. Some OECD countries have experienced a decrease in the size of the annual immigrant inflow between 1980 and 1995. Over the same years, the number of immigrants per year has increased in several other OECD countries.[1] The percentage change of the annual immigrant inflow from 1980 to 1995 ranges between negative 42% (in Japan) and positive 48% (in Canada). For all destinations, such changes

[1] There has been a decrease in France, Japan, Netherlands, and the United Kingdom. There has been an increase in Belgium, Canada, Germany, Luxembourg, Norway, Switzerland, and the United States (OECD, 1997).

A. M. Mayda (✉)
Department of Economics and School of Foreign Service, Georgetown University, Washington, DC, USA
e-mail: amm223@georgetown.edu

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_3

are anything but monotonic (OECD, 1997). The variation in terms of origin countries is remarkable as well (OECD, 1997).

Several factors are likely to influence the size, origin, and destination of labor movements at each point in time and contribute to the variation observed in the data. However, very few empirical works in the literature have tried to understand what drives international migration, perhaps due to past unavailability of cross-country data.

In turn, international migration has recently received a great deal of attention in light of research showing its beneficial effects from an economic-development point of view. For example, the recent literature has pointed out repeatedly the potential of free migration to produce large benefits—most likely greater than the gains from liberalizing existing trade barriers (Rodrik, 2002). To fully understand these and other effects, it is important to identify the forces and constraints that shape international migration movements.

In this paper, I empirically investigate the determinants—economic, geographic, cultural, and demographic—of bilateral immigration flows. My analysis is based on the predictions of a simple theoretical framework that focuses on both supply and demand factors. I use yearly data on immigrant inflows into fourteen OECD countries by country of origin, between 1980 and 1995. The source of this data is the International Migration Statistics for OECD countries (OECD, 1997), based on the OECD's Continuous Reporting System on Migration (SOPEMI).

My paper is related to a vast literature on the determinants of migration. Clark et al. (2007) and Karemera et al. (2000) both focus on the fundamentals explaining immigrant inflows into the United States by country of origin in the last decades. Other papers in the literature that analyze the determinants of migration to the U.S. are Borjas (1987) and Borjas and Bratsberg (1996). Hatton (2005) investigates trends in UK net migration in the last decades. Finally, Helliwell (1998) sheds light on factors affecting labor movements in his investigation of the magnitude of immigration border effects, using data on Canadian interprovincial, US interstate, and US-Canada cross-border immigration.

This paper makes three contributions to the literature. First, my analysis puts greater emphasis than previous works on the demand side of international migration, namely destination countries' migration policies. This change of perspective is important, given restrictive immigration policies in the vast majority of host countries. Second, my work is the first one I am aware of to use the OECD (1997) data on international migration to

systematically investigate the drivers of international flows of migrants. Previous works have either used country cross-sections (Borjas 1987, Yang 1995), or have focused on a single destination country over time (Borjas & Bratsberg, 1996; Brücker et al., 2003; Clark et al., 2007; Karemera et al., 2000) or a single origin country over time (Yang, 2003). By extending the focus of the analysis to a multitude of origin and destination countries and taking advantage of both the time-series and cross-country variation in the data, I can test the robustness and broader validity of the results found in earlier works.[2] Third, this paper carefully reviews and proposes solutions to various econometric issues that arise in the estimation, such as endogeneity and reverse causality. These econometric complications have not all been addressed in the previous literature.[3] Once I deal with them (e.g., by controlling for destination and origin countries' fixed effects and for year effects), my analysis both delivers estimates broadly consistent with the predictions of the international migration model and generates empirical puzzles.

According to the international migration model, pull and push factors have either similar-sized effects (with opposite signs), when migration quotas are not binding, or they both have no (or a small) effect on emigration rates, when migration quotas are binding. It is not clear, ex ante, which one of the two scenarios characterizes actual flows. Migration policies in the majority of destination countries are very restrictive, which should imply binding constraints on the number of migrants. On the other hand, even countries with binding official immigration quotas often accept unwanted (legal) immigration.[4] Restrictive immigration policies are often characterized by loopholes, that leave room for potential migrants to take advantage of economic incentives. For example, immigration to Western European countries still took place after the late 1970s, despite the official closed-door policy. Family-reunification and asylum-seekers policies can explain continuing migration inflows to Western Europe (Joppke, 1998).

[2] Since I began working on this paper, I have become aware of other related, but independent papers analyzing cross-country migration patterns: Alvarez-Plata et al. (2003); Pedersen et al. (2004) Pedersen et al. (2006). I discuss these very recent contributions to the literature below, in relation to the data I use and results I find.

[3] See also Alvarez-Plata et al. (2003) for an excellent discussion of the properties of different estimators of the determinants of migration flows.

[4] Notice that the data set I use only covers legal migration.

My empirical results are puzzling because they are in part consistent with the first scenario and in part with the second one. I find that pull factors—proxied by the per worker GDP in the destination country—significantly increase the size of emigration rates. This result is robust to changes in the specification of the empirical model. Both absolute and relative pull factors matter. That is, the emigration rate to a given destination is an increasing function of that country's per worker GDP and a decreasing function of the average per worker GDP of all the other host countries in the sample[5] (each weighted by the inverse of distance from the origin country). On the other hand, the impact of push factors—proxied by the per worker GDP in the origin country—is seldom negative as theory suggests would be the case with not-binding migration quotas and, when it is, the size of the effect is smaller than for pull factors and insignificant. Therefore my analysis finds evidence of an asymmetric impact of pull and push factors on emigration rates.[6]

The asymmetry is a familiar puzzle. For example, it has been documented in several works in the literature on internal migration (see, e.g., Hunt (2006) and the papers referenced in its footnote 4). Based on the existing literature, there might be numerous reasons for the asymmetry and possibly different ones operating across borders versus within country borders. At the national level, where migration quotas do not exist, Hunt (2006) provides an explanation of the asymmetry by breaking down data by age group: Origin region's unemployment rates (push factor) have an insignificant impact on migration flows because the insignificant effect for the young—who are not as sensitive to their own layoffs as the old—dominates the significant positive effect for the old. This explanation cannot be investigated at the international level because of data unavailability.

Another interpretation in the literature of the asymmetry is that migration quotas are effectively not binding but the impact of income

[5] Since the host countries in the sample receive a large fraction of immigrants in the world, it is not overly restrictive to focus on them. For example, according to the United Nations (2004), the list of leading host countries of international migrants in 2000—as measured by the percentage of the world's migrant stock in each of these countries—includes the United States (20%), Germany (4.2%), France (3.6%), Canada (3.3%), Australia (2.7%), United Kingdom (2.3%), Switzerland (1%), Japan (0.9%), and the Netherlands (0.9%) (see Table ii.3, p.30). These countries all belong to my sample.

[6] By asymmetry between pull and push factors, I mean that the coefficient of economic conditions in the source region does not have the expected sign, while the coefficient of economic conditions in the destination region is, as expected, positive and significant.

opportunities in the origin country is affected by poverty constraints, due to fixed costs of migration and credit-market imperfections (Lopez & Schiff, 1998; Yang, 2003). Since lower levels of per worker GDP in the source country both strengthen incentives to leave and make it more difficult to overcome poverty constraints, the net effect might be close to zero. In the empirical analysis I investigate this possibility and I find very weak evidence that my result on push factors is driven by poverty constraints in the origin country.

Yet an alternative explanation of my findings is that the asymmetric effect I estimate for pull and push factors is explained by the demand side of international migration—namely, migration policies—and not by the supply side as is often assumed in the previous literature. Changes in mean income opportunities in the destination country not only affect migrants' incentive to move there but also impact the political process behind the formation of migration policies. For example, in periods of economic booms, policymakers are better able to overcome political opposition to and accommodate increasing migration inflows.[7] *If migration quotas are binding*, the latter political-economy channel will be at work while the determinants on the supply side will have no (or a small) impact. This would explain the asymmetric effect I estimate for pull and push factors. While I do not investigate this interpretation directly,[8] I find evidence which is consistent with migration policy playing a constraining role. In the empirical analysis, I differentiate the effect of pull and push factors according to changes in destination countries' migration policy. I find that the effect of pull factors becomes more positive and the impact of push factors turns negative in those years when a host country's immigration laws become less restrictive. This is also true for the impact of other supply-side determinants such as geography and demographics (see below). In sum, my results suggest that migration quotas matter as they mitigate supply-side effects.[9]

[7] Hanson and Spilimbergo (2001) focus on US border enforcement and show that enforcement softens when the sectors that use illegal immigrants expand, which is evidence that migration policy is affected by changes in economic conditions in the destination country.

[8] This interpretation goes beyond the theoretical model in this paper, which assumes exogenous migration quotas. The empirical analysis of the endogenous determination of migration policy and its role in explaining the asymmetric effect of pull and push factors is outside the scope of this paper.

[9] This result is consistent with the findings in Hatton (2004) where emigration from Britain in the era of free migration (before 1914) is compared to emigration in 1950

My empirical analysis also finds that inequality in the source and host economies is related to the size of emigration rates as predicted by Borjas (1987) selection model. An increase in the origin country's relative inequality has a non-monotonic effect on the size of the emigration rate: The impact is estimated to be positive if there is positive selection, negative if there is negative selection. Among the variables affecting the costs of migration, distance between destination and origin countries appears to be the most important one: Its effect is negative, significant, and steady across specifications. On the other hand, there is no evidence that cultural variables related to each country pair play a significant role. Demographics—in particular, the share of the origin country's population who is young—shape bilateral flows as predicted by the theory. Since the effect of geography and demographics works through the supply side of the model, their impact should be even stronger when migration quotas are relaxed, which is what I find in the data.

Finally, I empirically investigate the importance of network effects. Since immigrants are likely to receive support from other immigrants from the same origin country already established in the host country, they will have an incentive to choose destinations with larger communities of fellow citizens. Network effects imply that bilateral migration flows are highly correlated over time, which is what the data shows. However, it is not clear how to interpret this result. While it is consistent with supply factors (i.e., network effects), it could also be driven by demand factors (family reunification policies, for example).

The rest of this paper is organized as follows. Section 2 presents a simple model of international migration. In Sect. 3 I describe the data sets used, while in Sect. 4 I discuss the estimating equations and some econometric issues that complicate the analysis. Finally, I present the main empirical results and additional results in, respectively, Sects. 5 and 6. Section 7 concludes.

onwards, when immigration policies were in place in the four main host countries of British migrants. The paper finds that, from the mid-1960s, the impact of economic and demographic forces "became less powerful as they were increasingly inhibited by immigration policies in the principal destination countries." (p.1).

2 Theoretical Framework

Both supply and demand factors affect international migration flows. Migrants' decisions to move, according to economic and non-economic incentives, shape the supply side of labor movements. The host country's immigration policy represents the demand side, namely the demand for immigrants in the destination country. The theoretical framework in this paper is closely related to the previous literature (Borjas, 1999; Clark et al., 2007), the main difference being the greater emphasis in my model on destination countries' immigration policy. I consider two countries: country 0, which is the origin of immigrant flows and country 1, which is the destination. I first focus on the supply side of immigration and look at the probability that an individual chosen randomly from the population of country 0 will migrate to country 1. In each country, wages are a function of the individual skill level (s_i). The wages that individual i receives in country 0 and would receive if he migrated to country 1 are respectively equal to $w_{0i} = \alpha_0 + \theta_0 \cdot s_i + \epsilon_{0i}$ and $w_{1i} = \alpha_1 + \theta_1 \cdot s_i + \epsilon_{1i}$, where the two disturbances have zero means over the origin country's population. In light of the empirical analysis below, based on aggregate data, it is helpful to rewrite individual i's wages in the two locations as a function of first and second moments of the income distributions (of the origin country's population) at home and abroad respectively:

$$w_{0i} = \mu_0 + v_{0i}, \text{where } v_{0i} \sim N(0, \sigma_0^2), \quad (1)$$

$$w_{1i} = \mu_1^0 + v_{1i}, \text{where } v_{1i} \sim N(0, \sigma_1^2), \quad (2)$$

where the correlation coefficient between v_{0i} and v_{1i} equals ρ_{01}, μ_0 equals $\alpha_0 + \theta_0 \cdot \bar{s}_0$ and μ_1^0 equals $\alpha_1 + \theta_1 \cdot \bar{s}_0$ ($\bar{s}_0$ is the mean skill level of the origin country's population).

Notice that μ_1^0, which is equal to the mean wage of the origin country's population if it all migrated to country 1, is different from $\mu_1 = \alpha_1 + \theta_1 \cdot \bar{s}_1$, which is equal to the mean wage of the destination country's population in country 1 ($\bar{s}_1$ represents the mean skill level of the destination country's population). This point will be relevant in one of the robustness checks in the empirical analysis.

I assume that each individual has Cobb Douglas preferences for the two goods produced in the world (x_A and x_B), which implies an indirect utility (function) from having an income y given by $v(p_A, p_B; y) = \bar{A}(p_A, p_B) \cdot y$. I assume that each country is a small open economy characterized by free trade with the rest of the world: [10] therefore goods' prices p_A and p_B, as well as $\bar{A}(p_A, p_B)$, are given and equal across countries.[11] An individual in country 0 will migrate to country 1 if the utility of moving is greater than the utility of staying at home that is, given the assumptions above, if the expected income in country 1 net of migration costs is greater than the expected income in country 0. Following the literature, I can define an index I_i that measures the net benefit of moving relative to staying at home for a risk-neutral individual i:

$$I_i = \eta_{01} \cdot w_{1i} - C_i - w_{0i}, \tag{3}$$

where η_{01} is the probability that the migrant from country 0 will be allowed to stay in country 1, and $C_i = \mu_C + v_i^C$, with $v_i^C \sim N(0, \sigma_C^2)$, represents the level of individual migration costs.[12] The correlation coefficients between v_i^C and (v_{0i}, v_{1i}) are equal to (ρ_{0C}, ρ_{1C}). The implicit assumption in (3) is that, if the migrant moves to but is not allowed to stay in the destination country, he still incurs the migration costs C_i and gives up the home wage w_{0i}. In other words, the individual migrates to the host country before knowing whether he will be able to stay (for a longer period of time) and gain the income w_{1i}.[13] Immigrants may not be able to stay in the host country because of quotas due to a restrictive immigration policy.

The probability that an individual chosen randomly from the population of the origin country will migrate from country 0 to country 1 therefore equals:

[10] Given free trade, what explains the difference in rates of return to labor across countries? The answer is that, besides free trade, the other conditions for factor-price-equalization are not satisfied: for example, if international productivity differences exist, then only *adjusted* factor-price-equalization holds.

[11] In the empirical analysis I adjust for international differences in goods' prices, using PPP income levels.

[12] I assume that each individual knows the wage levels w_{1i} and w_{0i} he would get in each location, the migration costs C_i, and the probability η_{01}.

[13] This assumption is consistent with the evidence that immigrants often arrive to a destination country with temporary tourist or student visas with the hope of being able to stay.

$$P = \Pr[I_i > 0] = \Pr[\eta_{01} \times (\mu_1^0 + v_{1i}) - (\mu_C + v_i^C) - (\mu_0 + v_{0i}) > 0], \quad (4)$$

which can be rewritten as $P = 1 - \Phi(z)$, where $z = -\frac{(\eta_{01} \cdot \mu_1^0 - \mu_0 - \mu_C)}{\sigma_v}$, σ_v is the standard deviation of $(\eta_{01} \cdot v_{1i} - v_{0i} - v_i^C)$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal. The probability in (4) is the *supply* emigration rate $\frac{I_{01}^S}{P_0}$, where I_{01}^S represents the size of the migration flow as determined by the supply side of the model and P_0 the population in the origin country.

Next, I assume that the destination country's immigration policy sets quantity constraints for immigrants coming from each origin country. Let I_{01}^D be the maximum number of migrants from country 0 allowed each year into country 1. These immigration quotas, which represent country 1 's demand for immigrants from country 0, may or may not be binding. Only in the latter case does the emigration rate we observe in the data ($\frac{I_{01}}{P_0}$) equals the *supply* emigration rate $\frac{I_{01}^S}{P_0}$ defined above. On the other hand, if quantity constraints are binding, $\frac{I_{01}}{P_0}$ will be less than $\frac{I_{01}^S}{P_0}$. In general, the emigration rate we observe in the data is equal to the minimum of $\frac{I_{01}^S}{P_0}$ and $\frac{I_{01}^D}{P_0}$, and is represented in Fig. 1 by the heavy lines, as a function of μ_1^0, μ_0 and μ_C. The figure assumes that quotas I_{01}^D are exogenous, which means that they are not affected by μ_1^0 nor by μ_0 nor by μ_C. This is a strong assumption that is questioned in the interpretation of the empirical results.

I assume that the probability η_{01} that the migrant from country 0 will be allowed to stay in country 1 is equal to $\min\{1, \frac{I_{01}^D}{I_{01}^S}\}$. It is then possible to derive testable predictions for the impact of μ_1^0, μ_0, and μ_C on the emigration rate from country 0 to country 1:

$$\frac{d(\frac{I_{01}}{P_0})}{d\mu_1^0} = \begin{cases} \frac{\phi(z)}{\sigma_v} > 0, \text{ if } \frac{I_{01}^S}{P_0} < \frac{I_{01}^D}{P_0}; \\ 0, \text{ if } \frac{I_{01}^S}{P_0} \geq \frac{I_{01}^D}{P_0} \end{cases} \quad (5)$$

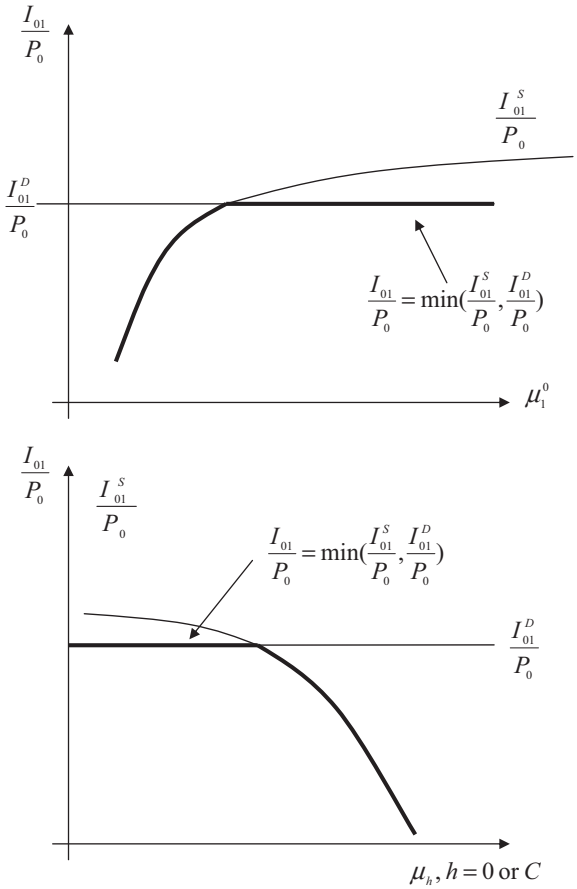

Fig. 1 The actual emigration rate as a function of mean income opportunities in the destination and origin country and of mean moving costs

$$\frac{d(\frac{I_{01}}{P_0})}{d\mu_h} = \begin{cases} -\frac{\phi(z)}{\sigma_v} < 0, \text{ if } \frac{I^S_{01}}{P_0} \le \frac{I^D_{01}}{P_0}; \\ 0, \text{ if } \frac{I^S_{01}}{P_0} > \frac{I^D_{01}}{P_0} \end{cases} \tag{6}$$

where $\phi(\cdot)$ is the density function of a standard normal and $h=0$, C. According to (5) pull effects (namely, improvements in the mean income opportunities in the destination country) are positive and strongest when restrictions are not binding neither ex-ante nor ex-post, they are positive but smaller in size when the quota is binding ex-post but not ex-ante and, finally, they are equal to zero in a quantity-constrained world. A parallel interpretation explains the comparative-static results in (6), which describes push effects (changes of μ_0, that is mean income opportunities in the origin country) and the impact of mean migration costs (changes of μ_C), according to the immigration-policy regime.

Thus, according to this simple model, pull and push factors have either similar-sized effects (with opposite signs), when quotas are not binding, or they both have no (or a small) effect on emigration rates, when quotas are binding. In the empirical analysis I will not be able to control for whether migration quotas are binding for a country pair in a given year (since I do not have data on I_{01}^{D}). Therefore I will estimate an average effect across country pairs with different degrees of restrictiveness. However, I will be able to use information on changes in I_{01}^{D}: I should find that pull (push) effects are more positive (negative) than average, for a given destination country, if that country's migration policy becomes less restrictive.[14]

Focusing for simplicity on the region where immigration quotas are not binding, it is straightforward to derive predictions for the impact of second moments of the income distributions (of the origin country's population) at home and abroad respectively. In particular, assuming that $\sigma_C=0$, we obtain the following expressions, where $k=\phi(z)(\sigma_1^2+\sigma_0^2-2\rho_{01}\sigma_0\sigma_1)^{-\frac{1}{2}}(-\frac{1}{\sigma_v^2})<0$ (Borjas, 1987):

$$\frac{d\left(\frac{I_{01}}{P_0}\right)}{d\sigma_1}=k\times(\mu_1^0-\mu_0-\mu_C)\cdot(\sigma_1-\rho_{01}\sigma_0), \tag{7}$$

[14] The reason is that, with higher I_{01}^{D}, the range of μ_1^0 (μ_0) for which the effect is strictly positive (negative) is wider (see Fig. 2).

$$\frac{d\left(\frac{I_{01}}{P_0}\right)}{d\sigma_0} = k \times (\mu_1^0 - \mu_0 - \mu_C) \times (\sigma_0 - \rho_{01}\sigma_1). \tag{8}$$

In my discussion I will assume that $(\mu_1^0 - \mu_0 - \mu_C) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The results in (7) and (8) imply that, if $\frac{\sigma_0}{\sigma_1} < 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_0}{\sigma_1}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$) will increase the emigration rate. Similarly, if $\frac{\sigma_0}{\sigma_1} > 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_1}{\sigma_0}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$) will decrease the emigration rate.

3 Data

In this paper, I merge data from an international migration panel with macroeconomic and other information on the origin and destination countries of immigrant flows. Data on immigration comes from the International Migration Statistics (IMS) data set for OECD countries (OECD, 1997), which provides information on bilateral immigrant flows based on the OECD's Continuous Reporting System on Migration (SOPEMI).[15] In particular, I use data on yearly immigrant inflows into fourteen OECD countries by country of origin, in the period 1980–1995. The IMS data only covers legal immigration; population registers and residence and work permits are the main sources of these statistics.[16] Based on this dataset, labor movements to the fourteen OECD countries appear to be both South-North and North-North flows. The sample includes seventy-nine origin countries with per worker GDP levels ranging from

[15] Alvarez-Plata et al. (2003) and Pedersen et al. (2004) use different international-migration data sets: the former paper uses the Eurostat Labor Force Survey which covers all destination countries within the EU-15 over nine years; the latter paper uses a dataset constructed by the authors after contacting the statistical bureaus in 27 selected destination countries (this data set covers the years between 1990 and 2000).

[16] Although the migration data is not perfectly comparable across OECD countries (some countries in the OECD (1997) data set define immigrants based on country of birth, while others based on citizenship), it is reasonable to think that changes over time can be compared.

approximately $1000 to $55,000 (PPP-adjusted) on average in the period considered.

The quality of the IMS data is high even though the coverage is not complete. The data set is supposed to cover immigrant inflows into each of the fourteen destination countries from all over the world. However, the sum by country of origin of the IMS numbers is not equal to 100% of the total flow into each destination country. The percentage of the total immigrant inflow covered by the disaggregate data ranges between 45% (Belgium) and 84% (United States). Put differently, the data set includes zero flows in correspondence of some country pairs (immigrant inflows from Italy to the United States, for example): some of these observations are likely to correspond to very small flows rather than zero flows. If very small flows are recorded as zeros in the disaggregate data set, there will be a discrepancy between total flows and the sum of flows by origin country. In the empirical analysis I will keep zero-flows observations in the data set and will investigate the robustness of my results to using a Tobit model.

Summary statistics and data sources for the other regressors used in the empirical model are documented in Appendix. Data on macroeconomic variables comes from various sources: the 2001 World Development Indicators data set (World Bank 2001) and the Penn World Tables (versions 5.6 and 6.1). Geographic and cultural information, such as on great-circle distance,[17] land border, common language, and colonial ties, comes from Glick and Rose's (2002) data set on gravity-model variables. I also use statistics on the average number of schooling years in the total population of destination and origin countries (over age 15) from Barro and Lee's (2000) data set. Data on Gini coefficients of destination and origin countries, used to construct the origin country's relative inequality variable, comes from Deininger and Squire (1996) data set (I only use so-called high-quality observations).[18] Finally, information on origin countries' share of young population comes from the United Nations.

Figure 2 shows that many destination countries in the sample are characterized by substantial volatility of immigrant inflows year after year. An important cause of variation over time in the number of immigrants to a given destination country is changes in that country's migration policy. For example, the United States' graph in Fig. 2 displays a peak around the

[17] Distance is calculated with the great circle formula using each capital city's latitude and longitude data.

[18] I linearly extrapolate data on schooling years and Gini coefficients for the years in which it is not available, based on the values for other years for the same country.

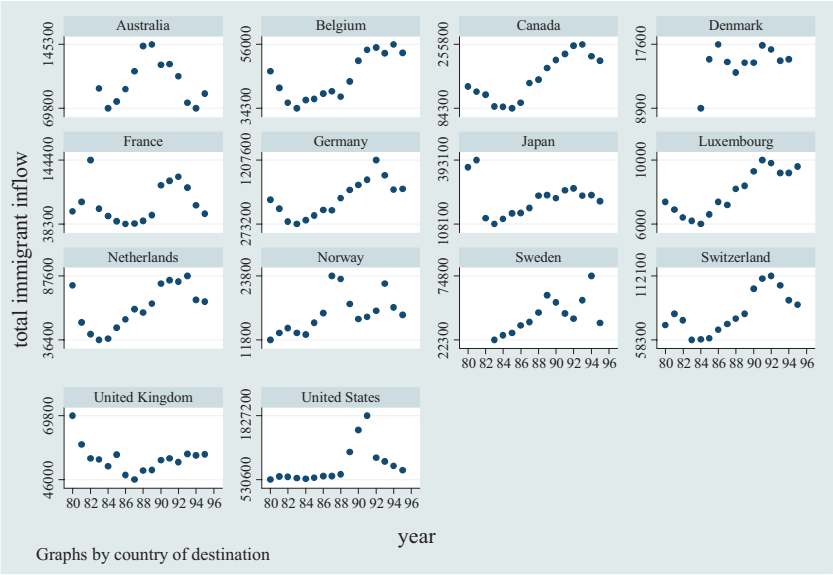

Fig. 2 Total immigrant inflow by destination country

year 1990. This is not surprising given that an amnesty law, the Immigration Reform and Control Act, was passed in 1986 and put in effect in the following years, with the bulk of the legalizations taking place in 1989–1991. The graph for Japan, on the other hand, displays a sudden decrease in the total immigrant inflow around the year 1982, which is when the Immigration Control and Refugee Recognition Act was passed. A separate Appendix to this paper documents the main characteristics of the migration policies of the destination countries in the sample and the timing (after 1980) of changes in their legislations (Mayda and Patel, 2004). A data set of destination countries' migration-policy changes, between 1980 and 1995, was constructed on the basis of the information in this Appendix and used in the empirical analysis.[19]

[19] In particular, the information in the Appendix (and in the background papers listed in the References) was used to identify: first, the timing of immigration-policy changes taking place in each destination country (the years in which migration policy laws were passed or enforced); second, the direction of the change in the case of substantial changes (loosening vs. tightening), based on a qualitative assessment of the laws (we mainly focused on aspects of migration policies related to the size of immigration flows, as opposed to, for example, issues of citizenship).

4 Empirical Model

According to the theoretical framework in Sect. 2, the estimating equation should include the emigration rate as the dependent variable and, among the explanatory variables, the mean wage of the origin country's population in, respectively, the origin and destination countries. As approximations for the latter two variables, I use the (log) level of per worker GDP, PPP-adjusted (constant 1996 international dollars) in the two countries.[20] Based on the theoretical model, I expect pull and push effects to be, respectively, positive and negative on average, if migration quotas are not binding, and both zero (or small) otherwise.

Another determinant of bilateral immigration flows implied by the model of Sect. 2 is the physical distance between the two locations, which affects migration costs C_i. The further away the two countries are, the higher the monetary travel costs for the initial move, as well as for visits back home. Remote destinations may also discourage migration because they require longer travel time and thus higher foregone earnings. Another explanation as to why distance may negatively affect migration is that it is more costly to acquire information ex-ante about far-away countries. Besides distance, I introduce additional variables that affect the level of migration costs C_i. A common land border is likely to encourage migration flows, since land travel is usually less expensive than air travel. Linguistic and cultural similarity are also likely to reduce the magnitude of migration costs, for example by improving the transferability of individual skill from one place to the other. Past colonial relationships should increase emigration rates, to the extent that they translate into similar institutions and stronger political ties between the two countries, thus decreasing the level of migration costs C_i.

Finally, I introduce the share of the origin country's population who is young (between 15 and 29 years old) as a demographic determinant of migration flows. Consider an extension of the basic model in Sect. 2 to a multi-period setting. In this set-up, the individual cares not only about current wage differentials net of moving costs, but about future ones too. This implies that a potential migrant from country 0 will have a bigger incentive to migrate the younger he is, as the present discounted value of net benefits will be higher the longer the remaining work life time is (for

[20] Unfortunately, wage data cannot be used because wage income series are not available for all countries (especially origin ones) in the sample. Since per worker GDP is not a direct measure of the mean wage of the origin country's population at home and abroad, I run robustness checks to test whether it is a good proxy for it.

positive I_i in each year). We would then expect the share of the young population in the origin country to positively affect the emigration rate out of that country.

In a cross-country analysis, such as in this paper, unobserved country-specific effects could result in biased estimates. For example, the estimate of the coefficient on the destination country's per worker GDP may be positive. Based on this result, it is not clear whether immigrants go to countries with higher wages or, alternatively, whether countries with higher wages have other characteristics that attract immigrants. Along the same lines, a negative coefficient on income at home leaves open the question of whether immigrants leave countries with lower wages or, alternatively, whether countries with lower wages have certain features that push immigrants to leave. To (partly) get around this problem, I exploit the panel structure of the data set and I introduce dummy variables for both destination and origin countries. This allows me to control for unobserved country-specific effects which are additive and time-invariant.[21] All the regressions also have year effects, to account for common time shocks, and robust standard errors clustered by country pair, to address heteroscedasticity and allow for correlation over time of country-pair observations. Notice that destination countries' fixed effects also allow me to control for features of their immigration policy which are time-invariant and common across origin countries. In order to capture the effect of *changes* in destination countries' migration policies, I introduce two interaction terms of an indicator variable of such changes with pull and push factors, respectively. According to the theory, if the migration policy of a destination country becomes less restrictive, the effect of pull (push) factors should turn more positive (negative).

The basic empirical specification thus looks as follows:

$$\begin{aligned}\frac{flow_{ijt}}{P_{it}} = {} & \beta + \beta_0 pwgdp_{it-1} + \beta_1 pwgdp_{jt-1} + \beta_2 dist_{ij} + \beta_3 border_{ij} \\ & + \beta_4 comlang_{ij} + \beta_5 colony_{ij} + \beta_6 youngpop_{it-1} \\ & + \beta_7 pwgdp_{it-1} \times immigpol_{jt} + \beta_8 pwgdp_{jt-1} \times immigpol_{jt} \\ & + \delta_i I_i + \delta_j I_j + \delta_t I_t + \varepsilon_{ijt}\end{aligned} \tag{9}$$

[21] In one robustness check, I control for country-pair fixed effects. In all the other regressions, I include separate destination and origin countries' fixed effects.

where i is the origin country, j the destination country, and t time. $\frac{flow_{ijt}}{P_{it}}$ is the emigration rate from i to j at time t ($flow_{ijt}$ is the inflow into country j from country i at time t, P_{it} is the population of the origin country at time t). *pwgdp* is the (log) per worker GDP, PPP-adjusted (constant 1996 international dollars) and *dist* measures the (log) great-circle distance between the two countries. The variable *border* equals one if the two countries in the pair share a land border. *comlang* and *colony* are two dummy variables equal to one, respectively, if a common language is spoken in the two locations, and for pairs of countries which were, at some point in the past, in a colonial relationship. The variable *youngpop* is the share of the population in the origin country aged 15–29 years old. The variable *immigpol* increases by one (decreases by one) if in that year the destination country's immigration policy became less (more) restrictive, zero otherwise. In other words, a change in policy is modelled as leading to a lasting effect (i.e., in the year when the policy change occurred and in the following years). Finally, the basic empirical specification also includes destination and origin countries' fixed effects (I_j and I_i) and year effects (I_t). According to the model in Sect. 2, I expect that $\beta_0 \leq 0$, $\beta_1 \geq 0$, $\beta_2 \leq 0$, $\beta_3 \geq 0$, $\beta_4 \geq 0$, $\beta_5 \geq 0$, $\beta_6 \geq 0$, $\beta_7 < 0$, and $\beta_8 > 0$.

An econometric complication is the possibility of reverse causality and, more in general, of endogeneity in the time-series dimension of the analysis. For example, the theoretical model in Sect. 2 predicts that, if migration quotas are not binding, better (worse) income opportunities in the destination (origin) country increase emigration rates. However, a positive β_1 (negative β_0) may just reflect causation in the opposite direction, that is the impact of immigrant flows on wages in host and source countries. After all, this channel is the main focus of analysis in many labor-economics papers (see Friedberg & Hunt, 1995 for a survey of this literature). More broadly, other time-variant third factors may drive contemporaneous wages and immigrant flows.

As for reverse causality, notice that it is likely to bias the estimates toward zero. The reason is that, if anything, immigrant inflows are likely to decrease wages in the destination country and outflows are likely to increase wages in the origin country. While the opposite signs are a theoretical possibility (e.g., in the economic-geography literature, because of economies of scale), the empirical evidence in the labor-economics

literature is that immigrant inflows have a negative or zero impact on the destination country's wages (Friedberg & Hunt, 1995; Borjas, 2003) and that immigrant outflows have a positive impact on the origin country's wages (Mishra, 2007).

Although reverse causality may not be an issue, it is still important to address endogeneity. Thus, I relate *current* emigration rates to *lagged* values of (log) per worker GDP, at home and abroad (and to *lagged* values of all the other time-varying regressors). While it is unrealistic to claim that wages at home and abroad are strictly exogenous, it is plausible to assume that they are predetermined, in the sense that immigrant inflows—and third factors in the error term—can only affect contemporaneous and future wages.[22]

5 Empirical Results

Table 1 presents the results from estimation of Eq. (9). The estimates show a systematic pattern, broadly consistent with the theoretical predictions of the international migration model. The analysis also generates empirical puzzles.

First, the emigration rate is positively related to the destination country's (log) per worker GDP. According to the estimate in regression (1), a 10% increase in the level of per worker GDP in the destination country increases emigration by 2.6 emigrants per 100,000 individuals of the origin country's population (significant at the 5% level). In other words, a 10% increase in the host country's per worker GDP implies a 20% increase in the emigration rate (as the mean of the dependent variable is, in regression (1), 13 emigrants per 100,000 individuals). This result would suggest that migration quotas are not binding on average across destination countries. However, the impact on the emigration rate of a change in the income opportunities at home is not consistent with this interpretation. Push effects are estimated to be insignificantly different from zero in Table 1 (and often of the wrong sign). One possibility is that, in practice, migration quotas are not binding, but push factors are zero due to the

[22] Strict exogeneity of an explanatory variable implies $E[X_{it}\varepsilon_{is}] = 0$, for $\forall s, t$, while predeterminacy implies $E[X_{it}\varepsilon_{is}] = 0$, for $\forall s > t$. In one of the following specifications, I also control for lagged values of the emigration rate, since if the emigration rate is autocorrelated, predeterminacy of the regressors does not guarantee consistency of the estimates.

Table 1 Determinants of bilateral immigrant flows

Equation	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Dependent variable	*Emigration rate*								
log per worker gdp (destination)	25.57	25.55	25.81	23.39	37.8	25.74	189.66	22.05	19.33
	10.79*	10.76*	10.79*	11.36*	15.56*	10.83*	76.66*	10.97*	14.55
log per worker gdp (origin)	2.05	1.78	4.82	6.01	-20.59	5.45	5.62	-7.97	9.74
	7.95	7.8	8.52	10.31	14.53	8.74	23.91	6.16	9.19
log distance	- 40.99	- 40.63	- 40.64	- 40.84	- 31.94	- 37.92	- 148.76		- 41.91
	9.51**	9.09**	9.09**	9.33**	10.51**	8.00**	4.46**		8.42**
Land border	- 28.14	- 36.94	- 36.93	- 38.88	- 17.48				
	19.69	23.24	23.28	24.7	27.68				
Common language		22.05	22.03	23.78	16.72				
		15.88	15.88	17.24	16.61				
Colony		3.02	2.9	1.96	11.1				
		16.91	16.94	18.03	18.75				
Share of young population (origin)			198.46	164.38	-10.2	204.45	573.14	124.97	223.43
			103.78+	104.73	112.51	106.00+	311.80+	82.85	109.42*
log yrs schooling (destination)				-29.76	-32.96				
				13.65*	23.83				
log yrs schooling (origin)				17.36	46.25				
				13.91	35.81				

(*continued*)

Table 1 (continued)

Equation	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Dependent variable	*Emigration rate*								
log capital per worker (destination)					-21.49				
					21.38				
log capital per worker (origin)					11.77				
					15.57				
Per worker gdp (destination)*immig policy change									17.63
									6.12**
per worker gdp (origin)*immig policy change									-3.24
									1.44*
log distance*immig policy change									-10.4
Share of young population (origin)*immig policy change									151.49
									48.73**
immig policy change									-109.19
									73.18
Number of observations	8010	8010	8010	7313	4103	8010	8010	8010	8010
R-squared	0.24	0.25	0.25	0.26	0.26	0.24		0.85	0.27

OLS estimates except for regression (7), which is a Tobit. All regressions include year effects. Destination and origin countries' dummy variables are included in specifications (1)–(7) and (9). Regression (8) includes country-pair fixed effects. Standard errors, clustered by country pairs, are presented under each estimated coefficient. Constant not shown. + significant at 10%; * significant at 5%; ** significant at 1%. See Appendix for data sources.

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population. per worker gdp is the level of per worker GDP, PPP-adjusted (constant 1996 international dollars), lagged by one year. distance is the great-circle distance. land border equals one if the destination and origin countries share a land border.

common language equals one if a common language is spoken in both host and origin countries. colony equals one for pairs of countries ever in a colonial relationship. share of young population (origin) is the share of the population in the origin country aged 15–29, lagged by one year. log yrs schooling is the log of the average schooling years in the population over age 15, lagged by one year. log capital per worker is non-residential capital stock per worker (1985 intl. prices), lagged by one year. immig policy change increases by one if in that year the host country's immigration policy became less restrictive, decreases by one if it became more restrictive, zero if there was no change.

effect of poverty constraints in the origin country. I will investigate this hypothesis in Table 2.

In regressions (1)–(3), Table 1, I also explore the role played by geographic (*log distance* and *land border*), cultural (*common language* and *colony*), and demographic (*share of young population* (*origin*)) determinants, respectively. The picture that emerges from my results is one in which geography and demographics are the most important among this set of drivers of migration flows. According to the estimate in column (1), doubling the great-circle distance between the source and host country decreases the number of emigrants by 41 per 100,000 individuals in the origin country (significant at the 1% level). On the other hand, a common land border does not appear to play a significant role. The impact of a common language, though of the right sign, is not statistically significant and, surprisingly, past colonial relationships do not appear to affect migration rates (this is true whether *common language* and *colony* are entered in

Table 2 Economic determinants more in detail

Equation	*1*	*2*	*3*	*4*	*5*	*6*
Dependent variable	*Emigration rate*					
log per worker gdp (destination)	6.64	26.42	25.51	17.1	23.02	37.78
	20.5	27.11	10.86*	12.1	11.15*	12.31**
log per worker gdp (origin)	4.12	4.5	75.7	8.87	6.73	1.68
	17.64	17.93	53.08	15.26	8.58	7.15
Square of log per worker gdp (origin)			-3.84			
			3.07			
Origin country's relative inequality	74.89	80.08				
	38.66+	44.06+				
Square of relative inequality	-28.11	-29.29				
	13.15*	14.51*				
Unemployment rate (destination)				-0.35		
				0.4		
Unemployment rate (origin)				0.96		
				0.91		
Multilateral pull					-9.33	
					5.05+	
Emigration rate (t-1)						0.66
						0.02**

(*continued*)

Table 2 (continued)

Equation	*1*	*2*	*3*	*4*	*5*	*6*
Dependent variable	*Emigration rate*					
log distance			-36.82	-30.43	-36.16	
			7.57**	8.28**	7.36**	
Common language			18.86	19.16	18.97	
			11.85	11.55+	11.86	
Share of young population (origin)			186.06	-35.59	195.65	
			105.09+	114.93	103.85+	
Constant	-143.56	-349.04	-327.12	-24.72	-4.54	-0.59
	274.49	333.02	234.08	220.43	186.64	0.24*
Number of observations	4028	3350	8010	5010	8010	6429
R-squared	0.18	0.18	0.25	0.23	0.25	

OLS estimates with year effects, except for regression (6) (see below). Destination and origin countries' dummy variables are included in each specification (except in regression (6)). Standard errors, clustered by country pairs, are presented under each estimated coefficient. + significant at 10%; * significant at 5%; ** significant at 1%. See the end of Table 1 for definitions of the main variables used.

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population.

multilateral pull gives, for each destination/origin country pair, the average of (log per worker gdp (destination)-log distance) over all the other destination countries. origin country's relative inequality gives a measure of the inequality in the origin country relative to the destination country (it equals the Gini coefficient in the origin country divided by the Gini coefficient in the destination country).

In regression (2), I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year. In Eq.(6), I include as regressors the emigration rate lagged by one, two, three, and four years (the coefficients on the latter three lags are not shown). Only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator). Column (6): Arellano-Bond test that average autocovariance in residuals of order 1 is 0: H0: no autocorrelation z =-55.05 Pr > z = 0.0000. Column (6): Arellano-Bond test that average autocovariance in residuals of order 2 is 0: H0: no autocorrelation z =-0.35 Pr > z = 0.7269. See Appendix for data sources.

the regression together or one at a time). Finally, the share of the origin country's population who is young has a positive and significant impact on emigration rates. A ten percentage point increase in the origin country's 15–29 years old population raises the emigration rate by 20 emigrants per 100,000 individuals (regression (3)).

Next, I investigate whether per worker GDP (PPP-adjusted) of origin and destination countries is a good proxy for mean income opportunities of migrant workers at home and abroad. Per worker GDP is not a direct measure of wages of a potential migrant, since it depends on rates of return

to both capital and labor and on endowments of each factor. For example, a higher per worker GDP in the destination country does not necessarily mean better income opportunities on average for an immigrant worker, since it could be due to a higher capital-labor ratio or to a more skilled labor force in the destination country's population. To address this concern, I run a robustness check where I control for the mean skill level and per worker capital endowment in destination and origin countries (columns (4)–(5)).[23] I first control for the average schooling level in both countries in regression (4). I still estimate pull effects which are positive and significant (at the 1% level). The results on push effects are the same as in previous estimates as well. In line with the theoretical predictions, the average skill level in the population of the destination (origin) country has a negative (positive) impact on the emigration rate. In regression (5) I control for the per worker endowments of both skill and capital and find that their coefficients are of the right sign (although not significant). Most importantly, my prior findings on pull and push factors are robust.

In column (6), out of all the geographic, cultural, and demographic determinants, I only include the ones which are significant based on regressions (1)–(3), that is *log distance* and *share of young population* (*origin*). I find evidence consistent with my previous results. Using a specification with these variables, I test how robust the results are—in particular, in terms of the asymmetry between pull and push factors—to using a Tobit specification (regression (7)). The estimates are again in line with the picture based on OLS regressions but they are larger in magnitude.

In the next regression (column (8)) I only exploit the variation over time within country pairs, by introducing fixed effects for each combination of origin and destination countries.[24] These country-pairs dummy variables allow me to control for time-invariant features of the destination country's immigration policy which are specific for each origin country. The results from this specification confirm that push and pull factors have an asymmetric effect in terms of magnitudes and significance levels.[25]

[23] Since capital is assumed to be internationally mobile, there are no international differences in rates of return to capital.

[24] Therefore I do not include the regressors *log distance*, *land border*, *common language* , and *colony* since they are constant within country pairs and, therefore, would be perfectly collinear with the country-pair dummy variables.

[25] If country pairs differ in terms of out-migration and return migration rates, net migration flows can be very different from gross flows. Since out-migration and return migration are likely to characterize specific country pairs, they are partially accounted for by including country-pair fixed effects.

Next, I investigate the interaction between changes in destination countries' migration policies and, respectively, pull and push factors (column (9), Table 1). Consistent with the theoretical predictions, positive pull factors are bigger than average for a destination country whose migration policy becomes less restrictive. Setting aside the average effect, push factors turn negative and significant once migration restrictions are relaxed. The opposite is true when policy becomes more protectionist. In the same regression I also add the interaction of the indicator variable of changes in destination countries' migration policy with, respectively, *log distance* and *share of young population* (*origin*). I find that the effect of the latter two variables is more pronounced (more negative and more positive, respectively) when a host country's immigration laws turn less restrictive. The opposite is true when policy becomes more protectionist. Notice that I also include the linear effect of immigration policy changes, which is insignificant. Regression (9) represents the preferred specification of the model. It shows that migration restrictions matter by mitigating effects on the supply side of the model (pull and push factors, geography, and demographics).

6 Additional Results

In Table 2, I analyze economic determinants more in detail. First, I investigate the impact of the second moments of the income distributions in the origin and destination countries. According to the theory (formulas (7) and (8)), given low values of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$), if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will increase, while given high values of $\frac{\sigma_0}{\sigma_1}$, if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will decrease.[26] The

[26] I assume that ρ_{01} is sufficiently high ($\rho_{01} > \max\{\frac{\sigma_0}{\sigma_1}, \frac{\sigma_1}{\sigma_0}\}$). The motivation for this assumption is explained in Borjas (1987): "It seems plausible to argue that for non-Communist countries, ρ_{01} is likely to be positive and large. After all, profit-maximizing employers are likely to value the same factors in any market economy" (p.534). I also assume that $(\mu_1^0 - \mu_0 - \mu_C) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The motivation for the last assumption is that the data set mostly includes migration flows from lower to higher average-income countries: the average difference in per capita GDP levels of destination and origin countries is positive and

intuition for these results is straightforward. If income inequality in the origin country is lower than in the destination country ($\frac{\sigma_0}{\sigma_1} < 1$), there is positive selection of immigrants from country 0 to country 1: migrants are selected from the upper tail of the income distribution at home and end up in the upper tail of the income distribution abroad (in both cases, the relevant distribution is the origin country's population one). For example, consider potential migrants from Portugal to the United States. Given that income inequality is lower in Portugal than in the U.S., among Portuguese workers it is the better-off who have an incentive to migrate while those at the very low tail of the income curve have an incentive to stay. The reason is that the probability of both very high and very low incomes is higher in the U.S. than in Portugal. An increase in income inequality in Portugal will make the marginal individual (who is in the lower tail of the income distribution) relatively worse-off at home and will increase her incentive to leave. Similarly, if income is more dispersed at home than abroad ($\frac{\sigma_0}{\sigma_1} > 1$), then there is negative selection of immigrants from country 0 to country 1: migrants are selected from the lower tail of the income distribution at home and end up in the lower tail of the income distribution abroad. An example of this situation is migration from Brazil to the U.S., given that income inequality in the latter is lower than in the former.[27] An increase in income inequality in Brazil will lower the emigration rate because those who were not migrating beforehand, the better-off, will have even less incentive to do so afterwards. In order to test these predictions, I introduce in the estimating equation a measure of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$) both in linear and quadratic forms. As expected, I find that the coefficient on the linear term is positive and on the quadratic term is negative (both significant at conventional levels),

substantial (approximately $20,600). I also add a robustness check (regression (2), Table 2) where I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year.

[27] The Gini coefficient for Portugal was 36.76 in 1990, while in the U.S. it was 37.8. The Gini coefficient for Brazil was 61.76 in 1985, while in the U.S. it was 37.26 (Deininger and Squire 1996).

which is consistent with Borjas (1987) selection model (regressions (1)–(2), Table 2).[28]

The remaining specifications in Table 2 investigate empirically a few extensions of the theoretical framework of Sect. 2. First, it is possible to incorporate poverty constraints in the model, due to fixed costs of migration and credit market imperfections in the origin country. As Yang (2003) shows, these assumptions imply that the effect on emigration rates of income opportunities at home is non-monotonic, positive at very low levels of income and negative for higher levels. Accordingly, I extend the empirical model previously specified by introducing both a linear and a quadratic term in per worker GDP of the origin country. I find very weak evidence of poverty constraints in regression (3). The sign of the coefficients is consistent with the theory but the lack of significance of the estimates prevents me from reading too much support into them.[29] This result thus leaves open the question of why push and pull effects are different in size and, indirectly, lends support to the alternative hypothesis of binding (and endogenous) migration quotas.

Next, the theoretical model can be modified by taking into account uncertainty in finding a job in each place. This extension suggests using the unemployment rate (which is approximately equal to one minus the probability of finding a job) as a regressor in the estimating equation. My results in column (4) are not significant. In an additional extension (column (5)), I test whether workers choose among multiple destination countries. In the theoretical model, the choice is between the origin country and one particular destination country. In practice, however, potential migrants are likely to compare mean income opportunities in their origin country to those in the destination country considered *and in any other*

[28] I evaluate the effect of relative inequality over the relevant range of values. Based on the coefficient estimates in column (1), Table 2, the threshold value of relative inequality is approximately equal to 2.6642: if $\frac{\sigma_0}{\sigma_1}$ is below this value (which is the case almost always in my sample, based on the summary statistics in Appendix), an increase in $\frac{\sigma_0}{\sigma_1}$ raises the emigration rate. This is consistent with positive selection taking place.

[29] In contrast with my results—which are not significant—Hatton and Williamson (2003) and Pedersen et al. (2004) find evidence of an inverted U-shaped effect on emigration of the origin country's economic conditions.

host country. For each pair of source and host economies, I construct and control for a *multilateral pull* term which is an average of per worker GDP levels of all the other destination countries in the sample, each weighted by the inverse of distance from the origin country. Regression (5) shows that third-country effects shape bilateral migration flows as expected, given that the coefficient of the *multilateral pull* term is indeed negative and significant (at the 10% level).[30]

To conclude, I investigate the role of past migration flows to the destination country from the same origin country. Lagged emigration rates capture the impact of network effects, which are likely to reduce the cost C_i of migration. The introduction of the lagged emigration rate among the explanatory variables makes the model a dynamic one. I use Arellano and Bond's GMM estimator to deal with the incidental parameter problem that arises with fixed-effects estimation of such a dynamic equation.[31] Emigration rates show considerable inertia in regression (6), where the coefficient of the lagged emigration rate is 0.66 (significant at the 1% level).[32] However, outside the model of Sect. 2—which assumes exogenous migration quotas—it is unclear how to interpret this autocorrelation. While it is consistent with network effects on the supply side, it could also be driven by factors working on the demand side. In particular, through the latter channel, past migration flows can influence the emigration rate in two different ways: through family-reunification immigration policies and through political-economy factors (see, e.g., Goldin (1994) and Ortega (2005), where the votes of naturalized immigrants affect immigration-policy outcomes).

[30] The multilateral pull term places migrants' decision to move in a multi-country framework. It is inspired by the multilateral trade resistance term in Anderson and van Wincoop (2003) (even though mine is an atheoretical measure).

[31] The Arellano and Bond estimator transforms into a difference the initial equation to remove the country-pair fixed effect and produces an equation that can be estimated with instrumental variables using a generalized method-of-moments estimator. The instruments include the lagged values of the dependent variable starting from t-4-2 (since the regression includes, as regressors, the emigration rate lagged by one, two, three, and four years).

[32] Regression (6) includes, as regressors, the emigration rate lagged by one, two, three, and four years (the coefficients of the latter three lags are not shown in the table). The reason is that, only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator).

7 Conclusions

In this paper, I empirically investigate the determinants of international bilateral migration flows. This analysis both delivers estimates consistent with the predictions of the international migration model and generates empirical puzzles.

In particular, I find evidence that pull factors, that is income opportunities in the destination country, significantly increase the size of emigration rates. This result is very robust to changes in the specification of the empirical model. On the other hand, the sign of the impact of push factors—that is, per worker GDP in the origin country—is seldom negative and, when it is, the size of the effect is smaller than for pull factors and insignificant. Therefore the evidence uncovered by the estimates is mixed in terms of the migration-policy regime that characterizes, on average, the destination countries in the sample: Push effects suggest that migration quotas are more binding than pull effects do. A possible explanation of the asymmetry between push and pull factors is the role played by the demand side of the model, that is destination countries' migration policies. While the theoretical framework of Sect. 2 assumes that migration quotas are exogenous, in practice they are not. Indeed migration policies can be thought of as the outcome of a political-economy model in which voters' attitudes toward immigrants, interest-groups pressure, policy-makers preferences, and the institutional structure of government interact with each other and give rise to a final immigration-policy outcome (Rodrik, 1995; Facchini & Willmann, 2005; Mayda, 2006). Binding and endogenous migration quotas can explain the asymmetric effect I estimate for pull and push factors. While I do not investigate the endogenous determination of migration policy, I find evidence consistent with the constraining role played by migration policies. In the empirical analysis, I interact an indicator variable of changes in destination countries' migration policies with pull and push factors, respectively. I find that pull effects become more positive and push effects turn negative in those years when a host country's immigration laws become less restrictive.

Among the variables affecting the costs of migration, distance appears to be the most important one. Its effect is negative, significant, and steady across specifications. Demographics, in particular the share of the origin country's population who is young, represent a significant determinant of emigration rates as well. I find that the effect of both variables is more pronounced in those years when a host country's immigration laws become less restrictive. In sum, my results suggest that migration quotas matter: They mitigate supply-side effects, that is pull and push factors, geography, and demographics.

The investigation of the determinants of international migration leads to other interesting research questions. The framework I have used in this paper to study migration flows is related to the gravity model of trade, which is used to analyze bilateral *trade* flows across countries. As a matter of fact, I have used several variables that appear frequently in the trade gravity literature (*log distance*, *land border*, *common language*, and *colony*). A common framework of empirical analysis for trade and migration makes it possible to combine the study of these two dimensions of international integration.

To conclude, by taking advantage of both the time-series and cross-country variation in an annual panel data set, this paper makes progress in explaining the determinants of international migration flows and in providing a framework for future analyses of migration relative to other dimensions of globalization.

Acknowledgments I would like to thank Alberto Alesina, Elhanan Helpman, Dani Rodrik, Klaus Zimmermann, and three anonymous referees for many insightful comments. For helpful suggestions, I am also grateful to Richard Adams, Marcos Chamon, Giovanni Facchini, Bryan Graham, Louise Grogan, Russell Hillberry, Arik Levinson, Lindsay Lowell, Rod Ludema, Lant Pritchett, Maurice Schiff, Tara Watson, Jeffrey Williamson, and participants at the International Workshop at Harvard University, at the 2003 NEUDC Conference at Yale University, at the International Trade Commission Workshop, at the World Bank International Trade Seminar, and at the IZA Annual Migration Meeting. I am grateful for research assistance provided by Krishna Patel and Pramod Khadka. All errors remain mine.

Appendix

See Table 3.

Table 3 Summary statistics (1980–1995) and data sources

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Emigration rate	8010	13.2433	81.5410	0.0000	1568.9430
Per worker gdp (destination)	8010	40682	5895	25252	55361
Per worker gdp (origin)	8010	20061	14106	1027	55361
log distance	8010	8.1715	0.8694	5.0872	9.3836
Land border	8010	0.0268	0.1616	0	1
Common language	8010	0.1704	0.3760	0	1
Colony	8010	0.0385	0.1923	0	1
Share of young population (origin)	8010	0.2612	0.0303	0.1951	0.3152
Years schooling (destination)	4103	9.6403	1.3096	6.8370	11.8650
Years schooling (origin)	4103	7.0285	2.4659	2.7240	11.8650
Capital per worker (destination)	4103	36041	12167	16992	76733
Capital per worker (origin)	4103	19232	13290	822	48135
Unemployment rate (destination)	5010	6.7306	3.4646	0.5000	14.1000
Unemployment rate (origin)	5010	8.1476	5.2840	0.0800	27.6000
Origin country's relative inequality	4028	1.2123	0.3846	0.3861	2.6810

The emigration rate (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) is from the IMS data set (OECD, 1997). Per worker GDP, PPP-adjusted (constant 1996 international dollars) is from the Penn World Tables, version 6.1. Log distance, land border, common language, and colony (countries ever in a colonial relationship) are from Glick and Rose (2002). Years of schooling are from Barro and Lee (2000) data set. Capital per worker (Nonresidential Capital Stock per Worker, 1985 intl. prices) is from the Penn World Tables, version 5.6.

The share of young population (origin) is based on data from the United Nations. The unemployment rate is from the World Development Indicators (2001), World Bank. The origin country's relative inequality is based on data on Gini coefficients from Deininger and Squire (1996) data set (only high-quality observations were used). The data set on immigration policy changes was constructed by Mayda and Patel (2004). All time-varying variables (except the emigration rate) are lagged by one year.

Summary statistics for the emigration rate, per worker gdp (destination), per worker gdp (origin), log distance, land border, common language, colony, share of young population (origin) are based on the same observations as in regressions (1)–(3) and (6)–(9), Table 1. Summary statistics for years of schooling (destination), years of schooling (origin), capital per worker (destination), capital per worker (origin) are based on the same observations as in regression (5), Table 1. Summary statistics for unemployment rate (destination) and unemployment rate (origin) are based on the same observations as in regression (4), Table 2. Finally, summary statistics for the origin country's relative inequality are based on the same observations as in regression (1), Table 2.

References

Alvarez-Plata, P., Brucker, H., & Siliverstovs, B. (2003). Potential migration from Central and Eastern Europe into the EU-15—An update Report for the European Commission, DG Employment and Social Affairs.

Anderson, J. E., & Wincoop, E. V. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review, 93*(1), 170–192.

Barro, R., & Lee, J. (2000). International data on educational attainment. Data Set.

Borjas, G., & Bratsberg, B. (1996). Who leaves? The outmigration of the foreign-born. *Review of Economics and Statistics, 78*(1), 165–176.

Borjas, G. J. (1987). Self selection and the earnings of immigrants. *American Economic Review, 77*(4), 531–553.

Borjas, G. J. (1999). The economic analysis of immigration. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (chapter 28, pp. 1697–1760). The Netherlands: North-Holland Elsevier Science.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics, 118*(4), 1335–1374. Harvard University.

Brücker, H., Siliverstovs, B., & Trübswetter, P. (2003). International migration to Germany: Estimation of a time-series model and inference in panel cointegration. DIW Discussion paper No. 391.

Clark, X., Hatton, T. J., & Williamson, J. G. (2007). Explaining US immigration, 1971–1998. *Review of Economics and Statistics, 89*(2), 359–373.

Facchini, G., & Willmann, G. (2005). The political economy of international factor mobility. *Journal of International Economics, 67*(1), 201–219.

Friedberg, R., & Hunt, J. (1995). The impact of immigrants on host country wages, employment, and growth. *Journal of Economic Perspectives, 9*(2), 23–44.

Glick, R., & Rose, A. (2002). Does a currency union affect trade? The time-series evidence. *European Economic Review, 46*(6), 1125–1151.

Goldin, C. (1994). The political economy of immigration restriction in the United States, 1890 to 1921. In C. Goldin & G. Libecap (Eds.), *The regulated economy: A historical approach to political economy* (pp. 223–257). Chicago, IL: University of Chicago Press.

Hanson, G. H., & Spilimbergo, A. (2001). Political economy, sectoral shocks, and border enforcement. *Canadian Journal of Economics, 34*(3), 612–638.

Hatton, T. J. (2004). Emigration from the UK, 1870–1913 and 1950–1998. *European Review of Economic History, 8*(2), 149–171.

Hatton, T. J. (2005). Explaining trends in UK immigration. *Journal of Population Eco nomics, 18*(4), 719–740.

Hatton, T. J., & Williamson, J. G. (2003). Demographic and economic pressure on emigration out of Africa. *Scandinavian Journal of Economics, 105*(3), 465–482.

Helliwell, J. F. (1998). *How much do national borders matter*? (chapter 5, pp. 79–91) Washington, DC: Brookings Institution Press.

Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association, 4*(5), 1014–1037.

Joppke, L. (1998). Why liberal states accept unwanted immigration. *World Politics, 50*(2), 266–293.

Karemera, D., Oguledo, V. I., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics, 32*(13), 1745–1755.

Lopez, R., & Schiff, M. (1998). Migration and the skill composition of the labor force: The impact of trade liberalization in LDCs. *Canadian Journal of Economics, 31*(2), 318–336.

Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *Review of Economics and Statistics, 88*(3), 510–530.

Mayda, A. M., & Patel, K. (2004). OECD countries migration policies changes. Appendix to International migration: A panel data analysis of the determinants of bilateral flows, by Anna Maria Mayda. http://www.georgetown.edu/faculty/amm223/

Mishra, P. (2007). Emigration and wages in source countries: Evidence from Mexico. *Journal of Development Economics, 82*(1), 180–199.

Nations, U. (2004). *World economic and social survey 2004. International migration*. New York: United Nations.

OECD (1997). International migration statistics for OECD countries. Data set.

Ortega, F. (2005). Immigration policy and skill upgrading. *Journal of Public Economics, 89*(9–10):1841–1863.

Pedersen, P. J., Pytlikova, M., & Smith, N. (2004). Selection or network effects? Migration flows into 27 OECD countries, 1990–2000. IZA Discussion Paper 1104.

Pedersen, P. J., Pytlikova, M., & Smith, N. (2006). Migration into OECD countries 1990–2000. In Parson & Smeeding (Eds.), *Immigration and the transformation of Europe*. Cambridge: Cambridge University Press.

Rodrik, D. (1995). Political economy of trade policy. In G. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 3, chapter 28, pp. 1457–1494). The Netherlands: North Holland Elsevier Science.

Rodrik, D. (2002). Comments at the conference on "Immigration policy and the welfare state" In T. Boeri, G. H. Hanson, & B. McCormick (Eds.), *Immigration policy and the welfare system*. Oxford: Oxford University Press.

Yang, D. (2003). *Financing constraints, economic shocks, and international labor migration: Understanding the departure and return of Philippine overseas workers*. Harvard University, dissertation chapter.

A unified model of relative deprivation and risk-laden migration

Oded Stark

1. Introduction

In this chapter we integrate two strands of research on migration that hitherto were studied independently. Theory and empirics have it that a sense of relative deprivation can be bothersome enough to induce people to resort to migration. (The earliest conceptualization of this relationship is by Stark, 1984. The earliest empirical validations are by Stark and Taylor, 1989, 1991. Follow-up empirical support is provided by, among others, Quinn, 2006; Stark et al., 2009; Czaika, 2011; Basarir, 2012; Jagger et al.,

The original version of the chapter has been revised. A correction to this chapter can be found at https://doi.org/10.1007/978-3-030-89258-6_29

A significantly expanded version of this chapter, which includes measures of robustness and model simulations, is entitled "An integrated theory of relative deprivation and risk-laden migration." It is co-authored with Wiktor Budzinski. And it is scheduled to appear in Sauer, Robert M. (ed.), *World Scientific Handbook of Global Migration*. Singapore: World Scientific Publishing. (In press.)

Oded Stark
University of Bonn, Bonn, Germany

University of Warsaw, Warsaw, Poland
e-mail: ostark@uni-bonn.de

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_4

2012; Vernazza, 2013; Flippen, 2013; and Kafle et al., 2020.) Both research and casual observation have shown that migration can be a risky undertaking. More than 50 years ago, the risk of unemployment following migration featured prominently in the modeling of migration (Todaro, 1969). Many studies that followed emphasized that migration is risky (David, 1974; Katz and Stark, 1986; Taylor, 1986; Heitmueller, 2005; and Bryan et al., 2014 are examples of such studies). The received literature did not acknowledge, though, that experiencing relative deprivation at origin not only encourages people to want to leave; it also makes them more willing to accept the risks involved in migration. Moreover, higher relative deprivation at origin is matched by riskier acceptable migration options. Formalizing this link yields a new testable model of migration.

The migration that we consider in this chapter is an individual act. A branch of research on migration and risk has, however, taken a distinct path, viewing migration as a means of risk diversification when the unit of analysis is the family, and when migration is by a family member such that the migrant on the one hand, and the family members who stay behind on the other hand, insure each other. Indeed, at the heart of earlier research on migration and risk, in particular in studies by Stark and Levhari (1982), Stark and Bloom (1985), Katz and Stark (1986), Rosenzweig and Stark (1989), and Stark (1993), lies the perception that the very purpose of migration by a family member is to reduce the risks that family members face. The course taken in this chapter is different in that the unit of analysis is the individual.

The setting is as follows. At origin, the individual experiences relative deprivation when, on comparison with other individuals, he observes that his income is lower than a certain measure of their incomes, let's say their average income. If he migrates, the individual will have to bear the risk of ending up with poor earnings or unemployment. When we model the individual's preferences and motivation, we find that higher relative deprivation experienced at origin is matched by riskier acceptable migration options.

2. An integrated model of relative deprivation and risk-laden migration

Consider individual i, $i \in \{1,2,\ldots,n\}$, who is a member of a population that consists of n individuals whose incomes are $0 < y_1 < y_2 < \ldots < y_n$. The individual derives pleasure from income, and displeasure from relative deprivation. The individual's satisfaction from income is represented by an increasing function $f(y_i)$ defined on $[0,+\infty]$. Being aware of the mean income in his population, $\overline{y}$, individual i experiences relative deprivation, $RD(y_i,\overline{y})$, if $y_i < \overline{y}$. We measure this relative deprivation by the distance

from below the mean income: $RD(y_i, \overline{y}) = \max\{\overline{y} - y_i, 0\}$, as is done, for example, in Stark (2013) and Stark (2017). The individual's utility depends on income, and on relative deprivation. We thus set the utility function of the individual to take the form

$$U_i(y_i, \overline{y}) = (1-\alpha_i) f(y_i) - \alpha_i RD(y_i, \overline{y}), \tag{1}$$

so that when the individual experiences relative deprivation, his utility function takes the form

$$U_i(y_i, \overline{y}) = (1-\alpha_i) f(y_i) - \alpha_i (\overline{y} - y_i), \tag{1'}$$

$U_i : [0,\infty) \times [0,\infty) \to \mathbf{R}$. The coefficients $\alpha_i \in (0,1)$ and $(1-\alpha_i) \in (0,1)$ in (1), are the weights that individual i assigns to his distaste for relative deprivation and to his preference for income, respectively. In using in the utility function weights that add up to 1, the function has the characteristic that a weak taste for absolute wealth is correlated with a strong distaste for low relative wealth (and vice versa).[1] This assumption can be interpreted as assigning 100 percent of weight to the absolute wealth and the relative wealth components, permitting any ratio between these two terms in the preference specification.

The comparison space of the individual, namely the domain in which the individual's relative deprivation is formed, is the population at the individual's location. The individual considers migrating, aware that migration poses a risk. To model this risk, we proceed as follows.

With probability $p \in (0,1)$, the individual will find work at his destination, in which case his income there will be x_i. With probability $1-p$, the individual will fail to find work at his destination, in which case his income there will be 0. We thus refer to income at destination as a random variable, X, such that $P(X = x_i) = p$, and $P(X = 0) = 1 - p$.

We denote the mean income at destination by $\overline{x}$. We assume that $\overline{x}$ is given; the arrival of individual i does not affect that mean income; from the perspective of the destination economy, migration is relatively small. To enable us to highlight the roles played by relative deprivation and risk in the inclination to migrate, we assume as follows.

First, that the individual experiences relative deprivation at origin, namely that $y_i < \overline{y}$.

[1] This characterization will hold also if we were to make the weaker assumption that $u(x_i) = af(x_i) - bRD_i(\mathbf{x})$ where $a, b > 0$, and $\mathbf{x}$ is the vector of incomes at destination.

Second, on migrating, the individual is in close social proximity to the native workers whose incomes are higher than his. Observing first-hand the mean income of these workers causes him to feel relative deprivation. Thus, even if the individual finds rewarding employment at his destination, he will be subject there to relative deprivation, namely $x_i < \bar{x}$. The assumptions that the individual experiences relative deprivation both at origin and at destination, namely that $y_i < \bar{y}$ and that $x_i < \bar{x}$, respectively, render it unnecessary to consider migration aimed at replacing relative deprivation at origin with no relative deprivation at destination.

The expected utility of the individual upon migration is

$$\begin{aligned}\mathbf{E}(U_i(X,\bar{x})) &= (1-p)\,U_i(0,\bar{x}) + pU_i(x_i,\bar{x}) \\ &= -(1-p)\,\alpha_i\bar{x} + p[(1-\alpha_i)\,f(x_i) - \alpha_i(\bar{x}-x_i)]. \qquad (2)\end{aligned}$$

A feature of our measure of relative deprivation is that migration by individual i does not discourage other individuals to consider migrating because the departure of any individual whose income is lower than the mean income raises the mean income of the remaining individuals.

Setting the direct cost of migration at zero, the individual will seek to migrate if $\mathbf{E}(U_i(X,\bar{x})) > U_i(y_i,\bar{y})$.

We denote the difference between the expected utility if migrating, (2), and the utility at origin, (1'), by $F(y_i,\bar{y},x_i,p)$, $F:[0,\infty)^3\times(0,1)\to\mathbf{R}$. The function $F(\cdot)$ takes the form

$$\begin{aligned}F(y_i,\bar{y},x_i,p) &\equiv \mathbf{E}(U_i(X,\bar{x})) - U_i(y_i,\bar{y}) \\ &= (1-\alpha_i)\,[pf(x_i) - f(y_i)] - \alpha_i(\bar{x}-\bar{y}-px_i+y_i). \qquad (3)\end{aligned}$$

Drawing on (3), we ask how the willingness of an individual to engage in risk-laden migration changes when his relative deprivation at origin changes.

Claim 1. Keeping the individual's income at origin constant, a higher relative deprivation at origin is matched by higher willingness to resort to risk-laden migration.

Proof. From (3) it follows that $\dfrac{\partial F(y_i,\bar{y},x_i,p)}{\partial\bar{y}} = \alpha_i > 0$: higher relative deprivation at origin, brought about by higher mean income there, $\bar{y}$, will, other things remaining unchanged, render risk-laden migration more appealing. Q.E.D.

We next consider a more stringent configuration, asking whether upon an increase in relative deprivation at origin the individual will be indifferent between staying at origin and undertaking a *riskier* risk-laden migration.

We thus consider a setting in which

$$F(y_i, \bar{y}, x_i, p) = 0, \tag{4}$$

namely a setting in which the individual is indifferent between staying at origin and undertaking a risk-laden migration. Keeping $\bar{x}$ and y_i constant, we are interested in signing the relationship $\frac{dp}{d\bar{y}}$, that is, while satisfying (4), we seek to ascertain the impact of relative deprivation (experienced in relation to mean income at origin) on the critical value of the parameter p, bearing in mind that this parameter represents the degree of risk involved in migration.[2]

Claim 2. Let an individual be indifferent between migrating and staying at origin, namely let $F(y_i, \bar{y}, x_i, p) = 0$. Then, upon an increase in relative deprivation experienced at origin, the individual who previously was indifferent between staying at origin and pursuing risky migration will be indifferent between staying at origin and undertaking *riskier* risk-laden migration.

Proof. Applying the implicit function theorem to $F(y_i, \bar{y}, x_i, p) = 0$, we get

$$\frac{dp}{d\bar{y}} = -\frac{\frac{\partial F(y_i, \bar{y}, x_i, p)}{\partial \bar{y}}}{\frac{\partial F(y_i, \bar{y}, x_i, p)}{\partial p}} = -\frac{\alpha_i}{(1-\alpha_i) f(x_i) + \alpha_i x_i} < 0. \tag{5}$$

Q.E.D.

The intuition underlying Claim 2 is as follows. To begin with, the individual who experiences relative deprivation as determined by his income at origin, y_i, and by the mean income at origin, $\bar{y}$, is indifferent

[2] One simple way of seeing this representation of the parameter p is to write the coefficient of variation $CV(X) = \frac{Std(X)}{E(X)} = \frac{\sqrt{x_i^2(p-p^2)}}{px_i} = \sqrt{\frac{1-p}{p}}$, for which $\frac{\partial CV(X)}{\partial p} = -\frac{1}{2p^2}\sqrt{\frac{p}{1-p}} < 0.$

between staying at origin and migrating. Migration is characterized by risk (finding employment is not certain). When the mean income at origin increases whereas the individual's income at origin remains unchanged, the individual is subject to greater relative deprivation. The level of p which "keeps" him on the indifference curve $F(y_i, \bar{y}, x_i, p) = 0$ then decreases. This is tantamount to an increase in the level of risk in admissible migration. In sum: experiencing higher relative deprivation at origin renders a previously unacceptable risk-laden migration acceptable.

3. Conclusion

In research on the causes of migration, a standard approach has been to focus on a given variable and attribute the urge to migrate to the expectation of gain in terms of this variable. An obvious example is wage earnings. In this chapter we depart from the direct link approach, exploring instead a crossover link. Another convention in migration research has been to differentiate between origin-based push causes and destination-based pull causes. In this chapter we deviate from this dichotomy, demonstrating that the line of demarcation between the two can be fuzzy rather than crisp sharp. And as we have seen, a result of these changes is the possibility of placing migration in a rich behavioral vein.

References

Basarir, Hasan (2012). "Is it all relative? The role of wealth in the migration decision: Evidence from Indonesia." Working paper, University of York.

Bryan, Gharad, Chowdhury, Shyamal, and Mobarak, Ahmed Mushfiq (2014). "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica* 82(5): 1671-1748.

Czaika, Mathias (2011). "Internal and international migration as response of double deprivation: Some evidence from India." Proceedings of the German Development Economics Conference, Berlin 2011, No. 21.

David, Paul A. (1974). "Fortune, risk, and the microeconomics of migration." In David, Paul A. and Reder, Melvin W. (eds.), *Nations and Households in Economic Growth: Essays in Honor* of *Moses Abramovitz*. New York: Academic Press.

Flippen, Chenoa (2013). "Relative deprivation and internal migration in the United States: A comparison of black and white men." *American Journal of Sociology* 118(5): 1161-1198.

Heitmueller, Axel (2005). "Unemployment benefits, risk aversion, and migration incentives." *Journal of Population Economics* 18(1): 93-112.

Jagger, Pamela, Shively, Gerald, and Arinaitwe, Arthur (2012). "Circular migration, small-scale logging, and household livelihoods in Uganda." *Population and Environment* 34(2): 235-256.

Kafle, Kashi, Benfica, Rui, and Winters, Paul (2020). "Does relative deprivation induce migration? Evidence from Sub-Saharan Africa." *American Journal of Agricultural Economics* 102: 999-1019.

Katz, Eliakim and Stark, Oded (1986). "Labor migration and risk aversion in less developed countries." *Journal of Labor Economics* 4(1): 134-149.

Quinn, Michael A. (2006). "Relative deprivation, wage differentials and Mexican migration." *Review of Development Economics* 10(1): 135-153.

Rosenzweig, Mark R. and Stark, Oded (1989). "Consumption smoothing, migration and marriage: Evidence from rural India." *Journal of Political Economy* 97(4): 905-926.

Stark, Oded (1984). "Rural-to-urban migration in LDCs: A relative deprivation approach." *Economic Development and Cultural Change* 32(3): 475-486.

Stark, Oded (1993). *The Migration of Labor*. Oxford and Cambridge, MA: Blackwell.

Stark, Oded (2013). "Stressful integration." *European Economic Review* 63: 1-9.

Stark, Oded (2017). "Migration when social preferences are ordinal: Steady-state population distribution and social welfare." *Economica* 84: 647-666.

Stark, Oded and Bloom, David (1985). "The new economics of labor migration." *American Economic Review* 75(2): 173-178.

Stark, Oded and Levhari, David (1982). "On migration and risk in LDCs." *Economic Development and Cultural Change* 31(1): 191-196.

Stark, Oded, Micevska, Maja, and Mycielski, Jerzy (2009). "Relative poverty as a determinant of migration: Evidence from Poland." *Economics Letters* 103(3): 119-122.

Stark, Oded and Taylor, J. Edward (1989). "Relative deprivation and international migration." *Demography* 26(1): 1-14.

Stark, Oded and Taylor, J. Edward (1991). "Migration incentives, migration types: The role of relative deprivation." *Economic Journal* 101: 1163-1178.

Taylor, J. Edward (1986). "Differential migration, networks, information and risk." In Stark, Oded (ed.), *Research in Human Capital and Development, Vol. 4, Migration, Human Capital, and Development*. Greenwich, CT: JAI Press.

Todaro, Michael P. (1969). "A model of labor migration and urban unemployment in less developed countries." *American Economic Review* 59(1): 138-148.

Vernazza, Daniel (2013). "Does absolute or relative income motivate migration?" London School of Economics, Mimeo.

High-Skilled Migration: Past, Present and Future

Alessandra Faggian

Migration is an expression of the human aspiration for dignity, safety and a better future. It is part of the social fabric, part of our very make-up as a human family.
—Ki-Moon (2013)

1 Introduction

Migrating is intrinsic in human nature. Campbell and Barone (2012), in their study on the origin of human migration, point out that proofs of human migration found in fossils date back at least 2 million years. The basic principle of migrating is clear, people move to "improve" their situation, but the meaning of "improve" has been changing over time and it is different across space. Initially individuals moved for survival reasons. As societies evolved, this sort of "forced" migration was gradually substituted with other forms of migration, where the scope was not strict survival, but

A. Faggian (✉)
Gran Sasso Science Institute (GSSI), L'Aquila, Italy
e-mail: alessandra.faggian@gssi.it

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_5

rather ensuring a better quality of life and living conditions. Although forced migration still exists in situations of conflict or extreme poverty, the majority of migration movements nowadays has a more "voluntary" component. Migrants move to get more opportunities, often linked to the labour market. High-skilled migrants are becoming an increasingly important part of migration flows. As Nathan (2014) highlights, in the decade 2000–2010, high-skilled (international) migrants (identified as individuals with a tertiary degree) in the OECD countries increased by 70%, as compared to 20% for the low-skilled. As Faggian et al. (2017) point out *labour markets have become increasingly globalized, which has created new opportunity for mobility, especially of high-skilled workers.* Acostamadiedo et al. (2020) expect a further increase in the number of migrants to the European Union by 2030. They propose four different scenarios and find an increase in 2030 of between 21% and 44% from the recorded average annual figure for the 2008–2017 period. The figure is much larger—between 83% and 208%—when restricted only to high-skilled labour migration. As such, migration is now a more "selective" process that has attracted a lot of attention of both academics and policy-makers. However, although the extensive, and growing, literature on the topic, much confusion still exists and more needs to be studied and understood. This short contribution aims at highlighting what we know on the topic, while also paving the way for possible future developments. Section "How Do We Define "High-Skilled" Migration?" starts by discussing what is meant by "high-skilled" migration; the other three brief sections sketch the main issues dealt with in the past, the improvements in the present and some possible future advancements.

2 How Do We Define "High-Skilled" Migration?

It is often the case that high-skilled migration is discussed without making explicit what is meant by "high-skilled." In reality, the way high-skilled migration is defined and measured in the literature is not univocal and there is no agreed concept, partly because the migrants themselves do not constitute a homogenous group (Salt, 1997). The three ways to identify high-skilled migrants found in the literature are based on: a. educational attainment; b. occupation; c. salary or income. Although the three alternatives are somewhat positively correlated, a contribution by Parsons et al. (2020) shows that their correlation is not as strong as previously thought. For instance, by looking at recent international immigrants to the USA,

using data from the American Community Survey for the year 2015, they find that only 8% of migrants can be defined as high-skilled if all three conditions, that is, a tertiary education, an occupation in the upper tier of the US Standard Occupational Classification and a salary of at least $100,000, are to be combined. If salary is relied upon in isolation, potentially 87% of otherwise classified high-skilled migrants may be omitted, if occupation is used this number goes down to 53%, but it is still very high. The most comprehensive measure is educational attainment (only 9% excluded). This reason, together with data availability, provides a strong justification in favour of education as a means to identify high-skilled migrants. While recognizing that also education has obvious limitations, education is therefore now the most common way to define high-skilled migrants (e.g. Pekkala et al., 2016; Parey et al., 2017; Basile et al., 2020; Crown et al., 2020a), although there are also examples of occupation-based definitions (e.g. Czaika & Parsons, 2017).

3 Past: What Do We Know About "High-Skilled" Migration?

Given that the initial study of migration goes back as far as Ravenstein in 1885, there is a lot we know about migration (including migration of highly skilled individuals), in particular with regards to two aspects:

(a) Migration **determinants**
(b) Migration **consequences**

Migration determinants include both individual and regional characteristics. Among the individual characteristics, some established facts include that age is negatively correlated with the probability of migrating, while education and skills are positively associated with it (Faggian et al., 2015). The reasons for the latter are many. Education helps in finding and processing information more efficiently (DaVanzo, 1983), it also implies a more "globalized" search for jobs, as educated individuals are more selective in picking their employment (Schwartz, 1976). More educated individuals also rely less on family and friends (DaVanzo, 1981) and are generally more *adaptable* to new places, in that they are more "receptive to change" (Levy & Wadycki, 1974) and extrovert (Crown et al., 2020b). In OECD countries, between 2000 and 2015, there has been a

continuous increase in the level of education of immigrants, and a simultaneous decline in the proportion of poorly educated ones.

As for regional characteristics, it is clear that individuals move to regions that offer not only better economic and labour market opportunities (Greenwood, 1985), but also better amenities (Graves, 1976, 1980). The balance between the two depends on the characteristics of the migrants themselves. For instance, younger migrants place more importance on labour market variables, while retirement migration is often linked to quality-of-life factors. The income of migrants affects their choices too, with richer migrants being able to give amenities a higher value.

As for the consequences of high-skilled migration, they are often supposed to be positive for the destinations, for example, in terms of increased productivity and innovation (Faggian & McCann, 2009), and negative for the origins, in the form of reduced growth and brain-drain (Beine et al., 2008; Wong & Yip, 1999). However, although maybe a bit counterintuitive, some argue that also origins might benefit from the emigration of highly skilled individuals, for instance, in the form of remittances, future return migration (also called brain-circulation) and the creation of networks between destinations and origins. Kanbur and Rapoport (2005) go an extra mile and point out that ex-ante emigration prospects could also foster the formation of human capital in the origin pushing more individuals to invest in their education. If not all of them emigrate, then those left behind represent an added-value for the origins.

4 Present: Improving Our Current Knowledge

In the past two decades, more and better data on migrants, both interregional and international, became available. This helped refining some of the previous work on both the determinants and consequences of migration. For instance, although a positive link between high-skilled migration and innovation was found in different contributions, all of them relied on patents to measure innovation, but, in time, it became clear that patents were a poor way to measure innovation especially for the sectors where high-skilled migrants were actually working.

As Faggian et al. (2017) state, *most highly-skilled migrants end up working in advanced service sectors where patents heavily underestimate innovation*. Pinate et al. (2021) extend the analysis of innovation, in the case of Italy, by using not only patents, but also other intellectual property rights (IPRs), that is, trademarks and design rights. As for the determinants,

better individual data means that it is possible to focus more on gender issues. The majority of migration studies in the past focused primarily on the experiences of male migrants (as head of the households) or did not distinguish between genders. This is a common flaw in the labour-related literature, but with better data it is possible to address the role of female migrants in more detail. This is of paramount importance because contributions on female migration are still rather scarce (e.g. Faggian et al., 2007) and results are mixed.

5 Future: Big Leaps with Big Data?

With the advent of so-called big data, there is a huge opportunity to further progress our knowledge of migration. Ascani et al. (2021), for instance, using data from Facebook, were able to study mobility patterns within and across local labour markets in Italy and the effect of the recent pandemic. The potential of these data is clear as they are high-frequency (three observations per day) and provide very large samples (about 2.5 million observations). However, with big potential come also possible big problems. First of all, it is easy to get lost in such a large amount of data and being tempted into data-driven, and not theory-driven, research. Second, there is the need to find appropriate tools of analysis (models) and even appropriate hardware and software. Third, there are issues of privacy and anonymity, which might prevent from getting detailed information on the characteristics of the migrants, such as their educational level or even age and gender, which are often the variables of interest.

References

Acostamadiedo, E., Sohst, R., Tjaden, J., Groenewold, G., & de Valk, H. (2020). *Assessing Immigration Scenarios for the European Union in 2030 – Relevant, Realistic and Reliable?* International Organization for Migration, Geneva, and the Netherlands Interdisciplinary Demographic Institute, the Hague.

Ascani, A., Faggian, A., Montresor, S., & Palma, A. (2021). Mobility in Times of Pandemics: Evidence on the Spread of COVID19 in Italy's Labour Market Areas. *Structural Change and Economic Dynamics, 58*, 444–454.

Basile, R., De Benedictis, L., Durban, M., Faggian, A., & Mínguez, R. (2020). The Impact of Immigration on the Internal Mobility of Natives and Foreign-Born Residents: Evidence from Italy. *Spatial Economic Analysis.* https://doi.org/10.1080/17421772.2020.1729997

Beine, M., Docquier, F., & Rapoport, H. (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *Economic Journal, 118*(528), 631–652.

Campbell, B. C., & Barone, L. (2012). Evolutionary Basis of Human Migration. In B. C. Crawford & M. H. Campbell (Eds.), *Causes and Consequences of Human Migration* (pp. 45–64). Cambridge: Cambridge University Press.

Crown, D., Faggian, A., & Corcoran, J. (2020a). High Skilled Immigration and the Occupational Choices of Native Workers: The Case of Australia. *Oxford Economic Papers, 72*(3), 585–605.

Crown, D., Gheasi, M., & Faggian, A. (2020b). Interregional Mobility and the Personality Traits of Migrants. *Papers in Regional Science, in press.* https://doi.org/10.1111/pirs.12516

Czaika, M., & Parsons, C. (2017). The Gravity of High-Skilled Migration Policies. *Demography, 54*, 603–630.

DaVanzo, J. (1983). Repeat Migration in the United States: Who Moves Back and Who Moves On? *Review of Economics and Statistics, 65*, 552–559.

DaVanzo, J. (1981). Microeconomic Approaches to Studying Migration Decisions. In R. W. Gardner (Ed.), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries.* Pergamon Press.

Faggian, A., Corcoran, J., & Partridge, M. (2015). Interregional Migration Analysis. In C. Karlsson, M. Andersson, & T. Norman (Eds.), *Handbook in the Research of Methods and Applications in Economic Geography* (pp. 468–490). Edward Elgar.

Faggian, A., & McCann, P. (2009). Human Capital, Graduate Migration and Innovation in British Regions. *Cambridge Journal of Economics, 33*(2), 317–333. https://doi.org/10.1093/cje/ben042

Faggian, A., McCann, P., & Sheppard, S. (2007). Some Evidence That Women Are More Mobile Than Men: Gender Differences in UK Graduate Migration Behavior. *Journal of Regional Science, 47*(3), 517–539.

Faggian, A., Rajbhandari, I., & Dotzel, K. (2017). The Interregional Migration of Human Capital and Its Regional Consequences: A Review. *Regional Studies, 51*(1), 128–143.

Graves, P. E. (1976). A Reexamination of Migration, Economic Opportunity and the Quality of Life. *Journal of Regional Science, 16*(1), 107–112.

Graves, P. E. (1980). Migration and Climate. *Journal of Regional Science, 20*, 227–237.

Greenwood, M. J. (1985). Human Migration: Theory, Models and Empirical Studies. *Journal of Regional Science, 25*(4), 521–544.

Kanbur, R., & Rapoport, H. (2005). Migration Selectivity and the Evolution of Spatial Inequality. *Journal of Economic Geography, 5*(1), 43–57.

Ki-Moon B. (2013). Secretary-General's remarks to High-Level Dialogue on International Migration and Development. Retrieved September 1, 2021, from https://www.un.org/sg/en/content/sg/statement/2013-10-03/secretary-generals-remarks-high-level-dialogue-international

Levy, M., & Wadycki, W. (1974). Education and the Decision to Migrate: An Econometric Analysis of Migration in Venezuela. *Econometrica, 42*(2), 377–388.

Nathan, M. (2014). The Wider Economic Impacts of High-Skilled Migrants: A Survey of the Literature for Receiving Countries. IZA *Journal of Migration, 3*(4).

Parey, M., Ruhose, J., Waldinger, F., & Netz, N. (2017). The Selection of High-Skilled Emigrants. *The Review of Economics and Statistics, 99*(5), 776–792.

Parsons, C. R., Rojon, S., Rose, L., & Samanani, L. (2020, September). *High Skilled Migration through the Lens of Policy Migration Studies, 8*(3), 279–306. https://doi.org/10.1093/migration/mny037

Pekkala Serr, S., Kerr, W., Çaglar, Ö., & Parsons, C. (2016). Global Talent Flows. *Journal of Economic Perspectives, 20*(4), 83–106.

Pinate, A., Faggian, A., Di Berardino, C., & Castaldi, C. (2021). *Migration and Innovation: Measuring the Effect by Origin and Skill Level on Patents, Trademarks, and Designs Across Italian Provinces.* GSSI Discussion Paper.

Ravenstein, E. G. (1885). The Laws of Migration. *Journal of the Statistical Society, 48*, 167–227.

Salt, J. (1997). *International Movements of the Highly Skilled*. OECD Report.

Schwartz, A. (1976). Migration, Age and Education. *Journal of Political Economy, 24*, 701–720.

Wong, K., & Yip, C. K. (1999). Education, Economic Growth, and Brain Drain. *Journal of Economic Dynamics and Control, 23*(5–6), 699–726.

From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China

AnnaLee Saxenian

Global labor markets are being transformed as the falling costs of transportation and communications facilitate greater mobility and as digital technologies support the formalization and long-distance exchange of large amounts of information. International migration, historically a one-way process, has become a reversible choice, particularly for those with scarce technical skills, and it is now possible to collaborate in real time,

---

AnnaLee Saxenian is dean and professor at the School of Information Management and Systems and professor in the Department of City and Regional Planning at the University of California, Berkeley. She is currently exploring how immigrant engineers and scientists are transferring technology entrepreneurship to regions in Asia. (*Studies in Comparative International Development, Summer 2005, Vol. 40, No. 2, pp. 35–61*)

A. Saxenian (✉)
School of Information, University of California, Berkeley, CA, USA
e-mail: anno@berkeley.edu; anno@ischool.berkeley.edu

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_6

even on complex tasks, with counterparts located at great distances. As a result, scientists and engineers from developing countries—once forced to choose between settling abroad and returning home to far less attractive professional opportunities—are contributing to their home economies while maintaining professional and economic ties in more technologically advanced economies. Some become "transnational" as they work, and even maintain residences and citizenship, in more than one nation.

The migration of talented youth from developing to advanced countries was viewed in the postwar decades as a "brain drain" that exacerbated international inequality by enriching already wealthy economies at the expense of their poor counterparts. According to a classic textbook on economic development,

> The people who migrate legally from poorer to richer lands are the very ones that Third World countries can least afford to lose, the highly educated and skilled. Since the great majority of these migrants move on a permanent basis, this perverse brain drain not only represents a loss of valuable human resources but could also prove to be a serious constraint on the future economic progress of Third World nations. (Todaro, 1985)

Data on these trends are hard to find, but the UN has estimated a total of 300,000 highly skilled emigrants from all developing countries to the West during the 1960s (Rapaport, 2002); the 1990 U.S. Census showed 2.5 million highly skilled immigrants, excluding students.

Much of the movement of skilled individuals from developing to advanced countries during the latter part of the twentieth century has involved migration to the United States, specifically Silicon Valley. The region's technology producers grew very rapidly from the 1970s through the 1990s, absorbing scientists and engineers voraciously and irrespective of national origin. Tens of thousands of immigrants from developing countries, who had initially come to the United States for graduate engineering education, accepted jobs in Silicon Valley rather than return to their home countries, where professional opportunities were limited.[1] By 2000, over half (53 percent) of Silicon Valley's scientists and engineers were foreign-born. Indian and Chinese immigrants alone accounted for over one-quarter of the region's scientists and engineers, or approximately

[1] NSF data show that over 95 percent of foreign-born engineering and science doctorate holders from India and China planned to stay in the United States after graduation.

20,000 Indian and 20,000 Chinese (5000 Taiwan- and 15,000 Mainland-born) engineers.[2]

This chapter argues that the same individuals who left their home countries for better lifestyles abroad are now reversing the brain drain, transforming it into "brain circulation" as they return home to establish business relationships or to start new companies while maintaining their social and professional ties to the United States. When foreign-educated venture capitalists invest in their home countries, they transfer first-hand knowledge of the financial institutions of the new economy to peripheral regions. These individuals, often among the earliest returnees, also typically serve as advisers to domestic policymakers who are anxious to promote technology growth. As experienced engineers and managers return home, either temporarily or permanently, they bring the worldviews and identities that grow out of their shared professional and educational experiences. These cross-regional technical communities have the potential to jump-start local entrepreneurship, and they succeed over the long term to the extent that they build alliances with technical professionals, businesses, and policymakers in their home countries.

The spread of venture capital financing provides a window into this process. In the early 1980s, returning immigrants began to transfer the Silicon Valley model of early-stage high-risk investing to Taiwan and Israel, locations that U.S. venture capitalists typically had neither interest in nor the ability to serve. Native-born investors provided the cultural and linguistic know-how needed to operate profitably in these markets. In addition to capital, they brought technical and operating experience, knowledge of new business models, and networks of contacts in the United States. Israel and Taiwan today boast the largest venture capital industries outside North America, and both have high rates of new firm formation and growth. Israel is now known for software and Internet firms like Mirablis (an instant-messaging program developer) and Checkpoint (security software); Taiwan has become a center of leading-edge personal computer (PC) and integrated circuit (IC) manufacturing with firms like Acer Technology Ventures (PCs and components) and TSMC (semiconductor foundry.) All have relied on returning scientists and engineers as well as a new breed of transnational venture investors.

[2] Indians accounted for 13 percent and Chinese for 14 percent of the region's engineers and scientists.

Building upon the experience of Israel and Taiwan, this chapter explores the developmental consequences of the heightened mobility of highly skilled workers in two of the world's largest developing economies: India and China. Specifically, it examines how Indian and Chinese immigrants to Silicon Valley are influencing economic development in their home countries directly, by transferring technology and know-how when they return home to work or start businesses, as well as indirectly, by influencing the formation of policy and other aspects of the institutional environment. Not surprisingly, this has proven to be a significantly easier process in Israel and Taiwan than in the complex political economies of China and India. Nonetheless, the long-term impacts of returning entrepreneurs and their communities may well be more far-reaching in the latter countries.

This chapter focuses on the creation of venture capital industries with close links to Silicon Valley—a process that entails extensive institutional change, particularly in domestic capital markets, and in turn has important consequences for the pattern of economic development. By 2004, venture capital and private equity firms were investing more than $1 billion annually in enterprises located in China and a comparable amount in India. While this is a fraction of the venture capital invested annually in the United States or even the amount of FDI in these economies, it supports indigenous entrepreneurship and has created an alternative, increasingly competitive, trajectory to the development opportunities provided by both the established domestic firms and the multinational corporations in these nations.

The first section of this chapter discusses the limits of traditional core-periphery understandings of the relation between developed and developing economies in an era of global labor mobility and "brain circulation," and particularly the failure to anticipate the development of independent technological capabilities in the periphery. The following section traces the transfer of the Silicon Valley model of venture capital to Taiwan during the 1980s by networks of U.S.-educated Chinese engineers. In this case the transfer of institutional and policy know-how was arguably as important as the later transfer of skills and technical knowledge. The third and fourth sections detail more recent processes of policy reform and institutional learning in China and India. The venture capital industries in these two countries have grown rapidly, with close connections to their Silicon Valley and U.S. counterparts, and in both countries, economic development in certain regions is characterized by high rates of entrepreneurship and experimentation. The chapter concludes with thoughts about the

extent to which it is possible to generalize from these cases of peripheral entrepreneurship to other late-developing economies.

1 Economic Development in an Era of Global Labor Mobility

Traditional accounts of economic development assume that new products and technologies emerge in industrialized nations that combine sophisticated skill and research capabilities with large, high-income markets—and that mass manufacturing is shifted to less costly locations once the product is standardized and the process stabilized. Success in this view builds on success in advanced economies, while peripheral economies remain followers. This divide is perpetuated by both the strategies of multinational corporations and the tendencies toward agglomeration created by the economics of increasing returns.

This model leaves little room for the development of independent technological capabilities in the periphery. At best, foreign investment from the core might contribute to the incremental mastery of manufacturing techniques and upgrading of local suppliers. Even the most successful newly industrializing countries are destined to remain imitators as long as leading-edge skill and technology reside in the corporate research labs and universities in the core. The primary route to development in the periphery, in this view, is the mobilization by the state, in conjunction with local banks and industry, of the resources to either develop or import the mass manufacturing capabilities that were perfected in the core.

Transformations in the world economy have undermined the power of this core-periphery model, however. The increasing mobility of highly skilled workers and information on the one hand, and the fragmentation of production in information and communication technology sectors on the other, provide unprecedented opportunities for formerly peripheral economies. Regions that missed the postwar economic boom, in particular, have provided fertile environments for a decentralized growth based on entrepreneurship and experimentation. The key actors in this process are neither policymakers nor multinational corporations in isolation, although both certainly play a role, but rather communities of technically skilled immigrants with work experience and connections to Silicon Valley and related technology centers.

U.S.-educated and trained engineers are increasingly transferring up-to-date technology and market information and helping to jump-start local entrepreneurship, allowing their home economies to participate in the information technology revolution. Because of their experience and professional networks, these cross-regional entrepreneurs can quickly identify promising new market opportunities, raise capital, build management teams, and establish partnerships with other specialist producers—even those located far away. The ease of communication and information exchange within ethnic professional networks accelerates learning about new sources of skill, technology, and capital as well as about potential collaborators. It also facilitates the timely responses that are essential in a highly competitive environment. This decentralized responsiveness is an advantage that few multinationals can claim.

This is not a one-way process. As recently as the 1970s, only large, established corporations had the resources and capabilities to grow internationally, and they did so primarily by establishing marketing offices or manufacturing branch plants overseas. Today the fragmentation of production and the falling costs of transportation and communication allow even small firms to build partnerships with foreign producers to tap overseas expertise, cost savings, and markets. Start-ups in Silicon Valley today are often global actors from their first day of operations; many raise capital, subcontract manufacturing or software development, and market their products or services outside the United States.

The scarce resource in this environment is the ability to locate foreign partners quickly and to manage complex business relationships and teamwork across cultural and linguistic barriers. This is particularly challenging in high-tech industries in which products, markets, and technologies are continually redefined—and where product cycles are often nine months or less. First-generation immigrants like the Chinese and Indian engineers in Silicon Valley who have the language, cultural, and technical skill to function well in the United States as well as in their home markets have a commanding advantage here. They have created institutions and social structures that enable even the smallest producers to locate and maintain mutually beneficial collaborations across long distances and that facilitate access to distant sources of capital, skill, and markets.

Late-developing economies typically face two major disadvantages: they are remote from the sources of leading-edge technology, and they are distant from developed markets and the interactions with users that are crucial for innovation (Hobday, 1995). Firms in peripheral locations use a

variety of mechanisms to overcome these disadvantages, from joint ventures and technology licensing to foreign investment and overseas acquisitions. However, a network of technologists with strong ties to global markets and the linguistic and cultural skills to work in their home country is arguably the most efficient and compelling way to overcome these limitations. Cross-regional entrepreneurs and their communities can facilitate the diffusion of technical and institutional know-how, provide access to potential customers and partners, and help to overcome reputational as well as informational trade barriers for isolated economies.

The increasing sophistication of information and communication technologies and the liberalization of global markets have accelerated this process. It is now quick, simple, and inexpensive to communicate internationally and to transfer information between distant locations. Information systems that facilitate the formalization of knowledge are dramatically expanding the volume as well as the variety of possible forms of information exchange. However, information technology alone cannot ensure successful coordination or efficient transfers of technical and institutional knowledge. Long-distance collaborations still depend heavily upon a shared social context and language that ensures mutual intelligibility between partners, particularly as speed and responsiveness are essential in today's technology competition.

Market liberalization has been equally important to the economic transformation of both China and India. However, the reduction of trade barriers and bureaucratic intervention alone does not create the institutional and social context, let alone the domain knowledge, required to sustain entrepreneurial success in global industries. Technology entrepreneurship remains highly localized even in the most advanced economies and it cannot be created by fiat, as evidenced by decades of failed attempts to "grow the next Silicon Valley." Efforts to jump-start entrepreneurship by mobilizing researchers, capital, and a modem infrastructure cannot replicate the shared language and trust of a technical community that permits open information exchange, collaboration, and learning (often by failure) alongside intense competition in places like Silicon Valley.

The new technology centers differ significantly from one another, and from Silicon Valley, in their technological sophistication as well as the specializations of local producers. Cross-regional entrepreneurs rarely compete head-on with established U.S. producers; instead they build on the skills and the technical and economic resources of their home countries. Israeli entrepreneurs, for example, have successfully applied the findings of

the nation's advanced military research to innovations in the Internet security and telecommunications arenas. Indian entrepreneurs, by contrast, recognized the opportunity to mobilize the thousands of underemployed English-speaking Indian engineers to provide software development services for American corporations. Returning entrepreneurs are ideally positioned to identify appropriate market niches, mobilize domestic skill and knowledge, connect to international markets, and work with domestic policymakers to identify and devise strategies to overcome obstacles to further growth.

These regions are each developing their own ecosystems for entrepreneurship, as well as close connections to technology and markets in the United States. The infrastructure for entrepreneurship is best-developed in Israel and Taiwan, where technologists have returned by the thousands since the 1980s and successfully transferred both U.S.-style venture capital and the Silicon Valley model of business focus and partnering. Both regions have also completed several entrepreneurial cycles in which successful entrepreneurs have reinvested their capital and contributed accumulated know-how and contacts to a subsequent generation of technology ventures, while also serving as role models. This cycle is central to establishing the relationships and decentralized information flow that support collective experimentation and learning in a regional economy. It does not guarantee the success of any individual firm, but it provides local producers with the capacity to collectively adapt and upgrade local capabilities.

The entrepreneurial ecosystem is still in its formative stages in the technology regions of India and China. These regions have seen important early entrepreneurial successes, and both have large technically skilled workforces willing to work very hard for relatively low wages. However, few U.S.-educated emigrants returned to either India or China during the 1980s and 1990s. The technology recession triggered an upsurge in cross-regional entrepreneurship that, along with the emergence of a second generation of successful start-ups, should significantly strengthen the entrepreneurial ecosystem of each. While it may be too early to define the trajectories of the regions in these large, complex political economies, the cross-regional entrepreneurs and their communities have pursued very different strategies. Returning Chinese entrepreneurs have focused primarily on developing products to serve the domestic market, while their Indian counterparts are oriented toward providing software and other services for export.

The dynamism of these technology regions is not reducible to cost advantages. Investors in India and China may have initially been motivated by the availability of low-cost skill, but the concentration of technology production has already generated rapidly rising wages and intensifying congestion in these regional economies. Engineering salaries in both Bangalore and Shanghai, for example, are now among the highest in their nations, yet new and established producers continue to cluster there rather than seeking lower-cost locations. The experience of Silicon Valley demonstrates that decentralized economies can flourish long after their labor cost advantages disappear as long as local investors and entrepreneurs are organized to collectively learn, innovate, and upgrade local capabilities—to create and recreate their regional advantage.

The contributions of an international technical community in transferring the institutions of technology entrepreneurship should not be confused with the broader role of a diaspora in the home country. The aggregate remittances, investments, or demonstration effects of a diaspora can affect an economy in a variety of different but largely limited ways. Transnational networks, however, are created by a small subset of highly educated professionals whose potential contributions to economic development are disproportionately significant. These transnational entrepreneurs are not typically drawn from the traditional economic or political elites of their home countries. Rather they are often the top engineering students from middle-class households whose access to education in the United States has landed them in a very different technological and institutional environment—one that they initially master and later transfer to their home countries.

Returning migrant communities are not replicating Silicon Valley around the world. Wide variations in national economic and political institutions, themselves the products of enormously varied histories and cultures, ensure distinctive and divergent economic trajectories. It is more appropriate to see the emerging regions as hybrids, combining elements of the Silicon Valley industrial system with inherited local institutions and resources. Returning entrepreneurs typically seek (with varying success) to transfer venture capital finance, merit-based advancement, and corporate transparency to economies with traditions of elite privilege, government control, and corruption. They seek to reproduce the team-based firm with minimal hierarchy and horizontal information flows in an environment dominated by family-run businesses or state-owned enterprises. The national institutions that support the Silicon Valley system—efficient and

well-developed capital markets, property rights, an independent judiciary, regulatory oversight, and sophisticated education systems, research institutions, and physical infrastructure—are rarely present in these peripheral economies.

Technology entrepreneurship is nonetheless a growing presence in each of these economies. Returning entrepreneurs have developed a variety of adaptations to the challenges created by conditions in their home countries. In India entrepreneurs rely on private telecommunications facilities and power supplies rather than on the nation's costly and unreliable infrastructure, while in China returning entrepreneurs have learned how to negotiate the complex rules that regulate private companies. Returning entrepreneurs also have the advantages of access to U.S. institutions: not only do they pursue graduate education in the United States, but many incorporate their businesses there, establish headquarters or research labs in Silicon Valley, seek venture capital and professional services as well as managerial and technical talent from the United States, and even raise money on U.S. capital markets.

At the same time, in all these countries, transnational entrepreneurs and their communities have devoted substantial time to efforts to transform domestic institutions by advising national governments on legal, regulatory, or capital market reforms, by working with regional governments on improving local infrastructure, universities, research, and training institutions, and by creating forums for information exchange and other forms of coordination among local producers. The outcome of these diverse strategies will differ from place to place—and while these regions may approximate the underlying principles of entrepreneurship-led growth, they do not seek to replicate the institutions or technological capabilities of Silicon Valley precisely.

A regional economic trajectory is shaped not only by local institutions but also by the range of technological and market opportunities available at the time it enters global markets. The most successful producers in Israel and Taiwan are those that have identified niches that allow them to differentiate and complement, rather than compete directly with, established producers in Silicon Valley—thus avoiding trade wars like those between U.S. and Japanese semiconductor firms in the 1980s, which reflected the old model, with vertically integrated "national champions" competing head-on for shares of a capital-intensive commodity business. And the fast-growing market for wireless communication in Asia has created opportunities for firms in China and India to contribute to the

direction of the technology and its applications—even if they do not define the leading edge of the technology. Over time, producers in developing regions build independent capabilities and define entirely new specializations and markets.

Organizational and institutional innovations will also likely emerge from these new centers of technology entrepreneurship, as they did from Japan in an earlier era. Entrepreneurship-led growth, with highly competitive and sophisticated small and medium-sized technology producers in high-skill regions connecting to and collaborating with counterparts elsewhere, is only one possible future for these formerly peripheral regions. They could forgo the opportunity to upgrade local skills and capabilities, and instead remain suppliers of low-cost labor to global (or domestic) corporations. China and India have the labor supply to do this for a relatively long time. However, many transnational entrepreneurs have maintained close ties to the technology and markets of Silicon Valley, and are constructing firms committed to an alternative, high-value-added trajectory.

2 Taiwan: Transferring the Silicon Valley Model

The Taiwanese economy emerged in the 1970s initially as a source of low-cost skill for labor-intensive calculator and later personal computer production, just as the old model predicts. The commitment to education provided the domestic skill base required for industrial upgrading, but at the same time networks of foreign-educated engineers worked with policymakers to develop local institutions that support entrepreneurial experimentation. Eventually networks of entrepreneurs and established firms defined specialized niches that allow them to focus and shift to higher-value-added activities without competing head-on with industry leaders. Taiwan became the world's most efficient center of IT manufacturing. While this process is at an earlier stage in China and India, and these are significantly larger and more complex political economies, it appears that both are following similar trajectories of upgrading.

Faced with the slow growth of the Science Park in Hsinchu, government officials sought alternative development strategies. K.T. Li, Taiwan's finance minister (who had visited Silicon Valley regularly in the 1970s to meet with Chinese engineers) was especially impressed by the U.S. venture capital industry and saw it as a potential missing link in Taiwan. In the early 1980s—long before it was fashionable elsewhere—he convinced the

Ministry of Finance of the need to provide funding for research-intensive production and promote the development of a public capital market. Once again Taiwan's policymakers studied the U.S. experience: they consulted investment professionals, organized collaborations with large U.S. banks to transfer financial and managerial expertise, and sent teams to Silicon Valley to be trained in managing a venture capital firm. Clark Su, chair of the Taipei Venture Capital Association, describes venture capital in Taiwan as a "pure transfer" from Silicon Valley.

Li spearheaded the legislation to create and regulate the venture capital industry and established a framework for enterprises to establish venture capital funds. This was challenging as the concept of venture capital was foreign to traditional Taiwanese practice, in which family members closely controlled all of a business's financial affairs. Under Li's guidance, the Ministry of Finance created significant tax incentives: 20 percent of the capital invested in strategic (technology-intensive) ventures by individual or corporate investors was tax-deductible for up to five years. Recognizing the challenge of raising capital from Taiwan's risk-averse financial and industrial communities, the government also provided substantial matching funds. The Ministry of Finance organized the initial "Seed Fund" with NT$800 million from the Executive Yuan Development Fund. This was so quickly allocated that the government committed a second fund of NT$1.6 billion.

Acer founded Taiwan's first venture capital firm, Multiventure Investment, in 1984 as a joint venture with Continental Engineering Group. Equally important, Li invited senior Chinese-American financiers to establish venture capital companies in Taiwan. Ta-lin Hsu, who had been a key senior policy adviser and STAG member since the 1970s, set up Hambrecht & Quist (H&Q) Asia Pacific in 1986. Hsu reports that it was not easy to raise the initial $50 million fund. In particular Li "twisted lots of arms" to raise $21 million (51 percent) from leading Taiwanese industrial groups such as Far East Textile, President Enterprises, and Mitac. The balance (49 percent) came from the government.[3]

The first general manager in H&Q Asia Pacific's Taipei office, Ding-Hua Hu, earned a doctorate in engineering at Princeton University in the 1970s and played a lead role in building Taiwan's semiconductor industry as the first general director of the Electronics Research and Service Organization (and previously an associate director of the Industrial

[3] Interview with Ta-lin Hsu, San Francisco, CA, June 1, 1997.

Technology Research Institute and a professor at Chiao Tung University). His career—and connections—underscore the extent to which the social and professional networks cut across university, government, and private sector (both financial and industrial) worlds in Taiwan and the United States. H&Q Asia Pacific's early investments included Acer, United Microelectronics Corporation (UMC), Microtek, and Tai Yan. The early successes of these investments made successive rounds of fundraising easier.

Two other U.S.-educated Overseas Chinese engineers, Peter Liu and Lip-Bu Tan, responded to Li's invitation as well, establishing Taiwan's second U.S.-style venture fund, the Walden International Investment Group (WIIG) as a branch of the San Francisco-based Walden Group in 1987. Both H&Q Asia Pacific and WIIG—along with Peter Liu's spin-off firm, WI Harper, remain leading investors in Taiwan's technology industries, and increasingly in Mainland China as well.

Once the early investments began to pay off, domestic IT firms created their own venture capital funds, including Acer, D-Link, Macronix, Mosel, Taiwan Semiconductor Manufacturing Company (TSMC), Silicon Ware, UMAX Data Systems, UMC, and Winbond. After that, even the old-line firms in traditional industries that had been reluctant also began investing in IT-related businesses.

The availability of venture capital transformed the Science Park from its originally envisioned role as an export processing zone into an open environment for the growth of indigenous technology firms. This was distinctive in Asia at a time when capital was available only to large corporations with ties to governments or to wealthy families. While throughout the 1980s Taiwan remained a low-value-added producer of electronics components, in the early 1990s—with the growth of entrepreneurship and accumulated production experience—local firms began to differentiate products on the basis of innovation and quality rather than simply low cost.

3 Building China's Venture Capital Industry

In the 1980s the Chinese government initiated creation of a domestic venture capital industry as a mechanism for supporting new technology ventures. But in contrast with Taiwan, where the venture capital industry took off in the 1990s, China's venture capital industry is constrained by the financial system and capital markets inherited from the planned economy. The China New Technology Venture Investment Corporation, which was established as a limited corporation in 1985 by the State Science

and Technology Council and the Ministry of Finance as the country's first experiment with venture capital, was declared bankrupt and closed by the People's Bank of China in 1997. This and several other early failures were an indication of structural problems, but did not stop local governments, universities and state-owned companies, and other organizations from setting up venture capital funds.

By 2000 there were approximately 160 domestic venture capital firms in operation in China, primarily located in Beijing, Shanghai, and Shenzhen. The underlying problem in many of these funds is that the government remains the primary source of capital, either directly or indirectly through university or state-owned firms. (There is ample private capital in China today but it is not being invested in venture capital funds.) This problem compromises the incentive for fund managers to make high-risk investments, particularly in private enterprises (White et al., 2002).

An investment in a government-owned company, by contrast, carries very little risk. Julie Yu Li, a partner in a venture capital firm started by a government-owned trading company in Shenzhen, summarizes the challenge: "I am supposed to invest in high-technology businesses, but my director once asked me if I could 'reduce the risk to zero'!" And she points out other limits of the current system: "I never expected my job would be this difficult. We have no way to identify entrepreneurs or to evaluate risks and returns and we must get approval from the president of the company to make any investments, which takes forever because it's so hard to get access to senior management."[4] A Silicon Valley-based entrepreneur who has advised the Chinese government on reforming the industry notes: "Venture capital fund managers in China have little at stake in the success of their ventures. If they are honest they will take no risk at all; if not they take advantage of the opportunity to make under-the-table deals with entrepreneurs."

The legal framework for venture capital is also not in place. For example, Company Law in China requires a minimal number of shareholders and level of investments (RMB 100,000 yuan per holder) that exceed the practice in the typical venture capital firm; likewise, it stipulates that the accumulated amount invested should not exceed 50 percent of the company's assets, which again arbitrarily limits the role of venture capitalists. Apparently these rules are being overlooked in practice (Xiao, 2002).

[4] Interview with Julie Yu Li, Shenzhen, China., January 11, 2001.

Some formal adjustments are being made. In 2001 the government issued a notice announcing that foreign companies would be allowed to set up wholly owned or Sino-foreign cooperative venture capital firms in China. Foreign venture capital firms are not, however, allowed to invest in securities, futures, or other financial markets, or in real estate and other industries that are not open to foreign investment. They are also not allowed to make loans or underwrite or invest with borrowed money. Other regulatory changes were targeted at the venture capital investments in technology firms, including reduction of minimum levels of capital invested, establishment of a preferential tax regime (10 percent) for venture capital investors, and recognition of the limited-liability partnership structure that is common to many venture capital funds in the West. Chinese Company Law also limits the amount of an enterprise's registered capital that can be granted for the contribution of intangible technology to a maximum of 20 percent. This limit has been abandoned in practice but not yet in regulations. However, the failure to specify restrictions on the qualifications for the general partners in a limited-liability partnership makes it likely that many more venture capital firms in China will fail to earn returns.

The problem is compounded by inexperience. Fund managers typically have little technical or business understanding of the industries they are investing in; they also lack procedures for objectively evaluating potential projects and ideas, and they have limited understanding of corporate governance. Potential entrepreneurs in China rarely know how to develop a viable business plan; they lack understanding of markets or management experience, and so they frequently adopt the model of the traditional Chinese family firm with husband and wife running the business rather than the professionally managed enterprise (which leads to struggles over ownership and control of the business). The managing director of the Beijing Venture Capital Company describes other challenges, including the information asymmetries between investors and fund managers and between venture capital firms and the companies they finance. In both cases the lack of transparency, objective performance measures, and external oversight creates incentives for concealing or falsifying information (Xiao, 2002). As a result, while the Chinese venture capital has financed thousands of high-tech enterprises, these funds have generated only

minimal returns.[5] A cynical interviewee claims that most venture capital in China is simply a new form of job creation for local governments.

The only real high-risk investors in China are the approximately fifty foreign venture capital firms that have a presence in the country—including many from Taiwan and Silicon Valley. However, these firms—including Japan's Softbank, Warburg Pincus, Intel VC, WI Harper, WIIG, H&Q Asia Pacific, Acer Technology Ventures, V2V, IDG, and Vertex (Singapore)—have invested quite cautiously in China. The overriding problem for these international investors is the lack of viable exit options, since access to the Chinese stock market has been impossible. Private enterprises in China have virtually no access to public capital markets because the two main trading boards, in Shanghai and Shenzhen, are dominated by former state-owned enterprises. Equally important for foreign investors, the RMB is not convertible, so there is no legal way to get earnings out of the country. International investors have consistently pushed for reforms to improve the environment for equity investments in China.

International venture capital has not only pushed hard for institutional reform but also financed some of China's most promising new technology enterprises. The successful firms typically have experienced senior management teams who are returning from the United States with sophisticated business plans, typically with the goal of serving the fast-growing China market. These "cross-Pacific start-ups" are typically headquartered in Silicon Valley, incorporated in the Cayman Islands or the Bahamas, and structured as U.S. Delaware-style corporations. Most have a permanent research and development capability in Silicon Valley as well as higher-level design or logistics capability in Taiwan.

These entrepreneurs and their investors are primarily Mainland-born or Taiwanese, with the first-hand knowledge of the language as well as culture and institutions that are essential to doing business in China. These returnees are often advantaged by networks of former classmates, friends, and family they can tap as they undertake the challenges of growing a firm in an environment that requires personal connections to get things done. Their investors typically bring knowledge not only of technology and management but also understanding of and experience in doing business in the Chinese market.

[5] Interview with Tang Hui-hao, Berkeley, CA, September 10, 2004.

The Chinese stock market was organized in the 1980s to provide capital for the expansion of state-controlled companies, and that bias continued during the 1990s as the exchanges were used to prop up failing state-owned enterprises. Even today, in spite of efforts at reform, the market remains poorly regulated and subject to price manipulation. The old governmental quota system, which gave each province an annual initial public offering (IPO) quota and ensured that state companies dominated the market listings regardless of their profitability, is being phased out. However, the top regulators and exchange officials for the China Securities and Regulatory Commission are government appointees, and promotion is based more on internal politics than on performance in managing the exchange.

The main trading boards have onerous listing requirements and complex regulatory procedures for approving public share offers that tend to favor state-backed firms. It took four years of intense lobbying for UFSoft to be listed on Shanghai's main board, in spite of its market leadership in accounting software, because of Chinese regulators' reluctance to approve privately owned firms. In 2000 only a handful of the thousand-plus companies listed on the Shanghai and Shenzhen exchanges were started as private companies.

Some Chinese policymakers are actively seeking to improve the environment for venture capital, and the past few years have seen almost continuous reform. The China Securities and Regulatory Commission in the late 1990s planned a second board in Shenzhen modeled after NASDAQ for high-risk, high-return companies. However, the anticipated opening was delayed indefinitely following the U.S. technology stock collapse in early 2001. An initial step toward setting up the second trading board was taken in 2004 when small and medium-sized firms were allowed to list on the Shenzhen exchange for the first time, and the expectation was that a second step would be taken in 2005 with a lowering of the listing requirements for these smaller capitalized companies. On the other hand, when a local government issues a list of the industries that venture capital will focus on in the coming years, it shows that China is often still trying to "grow the new economy using the tools of the planned economy" (Xiao, 2002).

Sustainable Start-ups

Acer Technology Ventures, the investment arm of Acer Computers, has played a central role in promoting cross-Pacific start-ups through investments in companies based in the United States and Asia. Its "IP Fund One" is a limited partnership incorporated in the Cayman Islands, devoting its $260 million to early-stage start-ups in the Internet protocol (enabling technology and solutions based on Internet platform) and intellectual property (software, integrated circuit (IC) design, etc.) fields. The limited partners include Acer affiliate companies (32 percent), Acer top management (6 percent), and institutional investors. With offices in Silicon Valley, Taipei, Shanghai, and Singapore, the firm markets itself as bringing "unique intangible value to portfolio companies with cross-regional business concepts." In particular, the firm provides "Cultural understanding to assist entrepreneurs in breaking new grounds cross-regionally in North America and the Asia Pacific region" (www.acervc.com). This assistance includes both communications and connections in these distant locations.

The sustainable cross-Pacific start-up, according to Acer's Ronald Chwang, is a technology-intensive business that combines Silicon Valley's new product and business vision, technology architecture, product marketing, and research and development coordination with China's research and development implementation, manufacturing and production logistics, and field engineering and local sales support. This leverages the management experience of Silicon Valley's overseas Chinese, China's low-cost engineering resources, standard and commercially available development tools, and the supply-chain manufacturing infrastructure of Greater China (Taiwan and Mainland). In this vision, products like mobile appliances or new semiconductor designs can rely on the U.S. market as an early drive and then be commercialized in both China and U.S. markets.

There remain substantial challenges to these cross-regional start-ups. Venture investment in China is still in its early stages: there has not yet been a complete cycle of investments and reinvestments in a second generation of entrepreneurs. Cross-regional firms also face significant difficulties coordinating distant activities, particularly in developing organizational synergy and persistent, consistent communication. Finally, they face the challenge that all technology firms in China face, of controlling and protecting their intellectual property (Chwang, 2003).

The Acorn Campus was established in Shanghai by a team of a half-dozen experienced overseas Chinese engineers—Taiwanese as well as

Mainland-born. General partner Wu-Fu Chen, for example, is one of Silicon Valley's most successful repeat entrepreneurs and a leader in the field of telecommunications and optics. The campus is an incubator in which they serve as "angel investors" and provide mentoring and connections, as well as space, for promising new ventures with Chinese founders. One of their recent investees, Newtone Communications, a telecom software firm, realized that its seed money of $500,000 would go much further in China than in the Bay Area. After moving to Shanghai, these returnees doubled their employment without increasing their budget.

This experience spurred the creation of a new Acorn Campus in Shanghai. In raising money for the Acorn Campus Asia Fund, the founders' mission is to "leverage the highest level of Silicon Valley entrepreneurial experiences to create, invest, and incubate high technology startups in China … and promote global leadership through Silicon Valley-Asia value chain partnerships" (www.acorncampus.com). Like WIIG, they address returning Silicon Valley Chinese entrepreneurs with substantial experience, and their focus is on semiconductor design, wireless infrastructure, and system and software development. They talk about accessing the best resources from different locations: research and development, new product development, and marketing in the United States; high-end logistics, design, and manufacturing in Taiwan; and low-cost engineering and manufacturing talent in China.

The power of the overseas community is most evident in the semiconductor industry, which originated in Silicon Valley and has been transferred by Chinese entrepreneurs first to Taiwan and then from Taiwan and Silicon Valley to China. In the words of WebEx CEO Min Zhu, "Silicon Valley is the technology leader and the center for real innovation because it supports the growth of start-ups. New firms cannot grow this fast in China or India. The most powerful model is a truly international company that combines the creative ideas and architectures that are developed in the United States with the ability to quickly implement them where skill is less expensive: both need to be scaleable."[6]

Of course, these are not one-way flows. While the Taiwanese IC industry initially grew out of talent and technology from the United States, producers like TSMC contributed indigenous innovations that in turn benefited the entire industry. It took follower UMC about a decade to reduce a process-technology gap of two or three generations to one

[6] Interview with Min Zhu, San Jose, CA, April 16, 2001.

generation (by the mid-1990s); today the firm is very close to the global leaders. China remains at a lower technological level than either Taiwan or the United States but now has in place the skill, capital, know-how, and connections to learn by doing. Over time the Chinese market will provide local IC designers and manufacturers the opportunity to experiment with, and ultimately innovate in, fields like wireless communications.

It is easy to overstate the scale of the domestic Chinese market in the short term. However, the internal demand for technology products remains relatively low; the domestic fabricators are serving foreign customers primarily. For example, Semiconductor International Corporation's early customers were overwhelmingly outside of China: according to the company's 2004 10K filing, in 2003, 36.7 percent of customers were located in North America, 26.7 percent in Taiwan, 12.5 percent in South Korea, 11.2 percent in Japan, 11 percent in Europe, and 1.9 percent in the rest of the Asia-Pacific region, including China. There is substantial room for growth of the domestic semiconductor market, but China will most likely continue to consume relatively low-end chips (the type used in watches, radios, cell phones, and other consumer electronics products) for the next five to ten years. IC manufacturing technology in China continues to lag behind Taiwan, and U.S. regulations on export of the most advanced manufacturing equipment to China will likely continue to slow the adoption of leading-edge process technologies.

And while research and development expenditures have increased significantly, China's companies, universities, and researchers remain relatively isolated from the market. As Frank He, president of a Dallas-based software company, puts it, "In China, research is too far from being applied to real life. For example, we can barely find one out of five hundred Ph.D. dissertations (e.g. in Computer Science) with research findings useful for commercialization. While in the United States, one dissertation out of one hundred can be turned into a real product" (Saxenian, 2003). McKinsey & Co. consultants in Shanghai predict that the large supply of low-cost engineering talent will allow China to grow more quickly as a center for semiconductor design than for advanced manufacturing, which requires sophisticated technology and management skills. Salaries for chip designers in China are about 20 percent of those in the United States, and the domestic market for IC design in China will reach an estimated $10 billion in 2010. According to Woetzel and Chen (2002), a senior IC design engineer with five years' experience or more earns a salary of US$14,000–$30,000 in China compared to $80,000–$150,000 in the

United States; likewise a junior design engineer in China earns $9000–$20,000 compared to $50,000–$100,000 in the United States.

The circulation of world-class engineering and entrepreneurial talent between the United States, Taiwan, and China is altering the economic trajectories of all three. The coming years pose a serious challenge to the adaptive capacities of Taiwan's industrial system. Experts claim that Taiwan has only ten years before China reaches technological parity in manufacturing of even the most advanced ICs and IT products (Clendenin, 2003). During this time, Taiwanese industry needs to establish new specializations, either as producers of higher-value-added software, content, and services or as manufacturers of distinctive systems and equipment that build on their existing strengths in designing, manufacturing, and recombining intermediate parts and components.

Semiconductor equipment maker Applied Materials predicts that by 2010 China will account for 7 percent of global semiconductor production, compared to 0.5 percent in 2000, and it will be a global center for low-cost IC manufacturing and regional distribution. Taiwan will, in this view, play a complementary role to China by transforming itself into a center for high-value-added IC design, productization, and advanced IC manufacturing. The United States (and Silicon Valley in particular) will remain the center for systems and chip architecture, capital investment, and production of IC manufacturing equipment.

The Taiwan-China connection will most likely continue to grow and deepen. Taiwan's 2001 IT investments of $2.6 billion in China are forecast to grow at compound annual rate of 32 percent and reach $10.6 billion in 2006; at that point Taiwanese IT investments in China are expected to surpass their investment in Taiwan (Chan et al., 2002), driven both by the rapid growth of China's domestic market for IT products and by the falling tariffs as a result of China's continued integration into the WTO.

The relationship between China and Taiwan, like that between the United States and Taiwan, is increasingly complementary rather than competitive, with Taiwan moving up the value chain to provide leading-edge manufacturing services and high-value-added design while China becomes a center of low-end, labor-intensive design and assembly-and-testing as well as non-leading-edge manufacturing.

4 Venture Capital and Entrepreneurship in India

Veteran Silicon Valley venture capitalist William H. "Bill" Draper III took U.S.-style early-stage venture investing to India in 1995. Soon after establishing a $55 million fund, Draper International, with a colleague in San Francisco and local partners in Bangalore and Bombay, he realized the many obstacles involved in working with entrepreneurs in India. The concept of venture capital was new in India, where the closely held, family-centered business model dominated both small firms and the large groups like Birla and Tata. It was very difficult to coordinate activities between India and Silicon Valley, and start-ups—with far more limited resources—faced all the same frustrations experienced by multinational investors, from the frequently corrupt and restrictive bureaucracy, to the lack of Western-style stock option plans to retain key personnel to the limits of basic infrastructure such as power and water and the very high cost of connectivity. In the words of Draper's partner, Robin Richards, in 1997:

> From a venture capitalist's perspective, it is 5–10 times slower to start up a company in India than in the U.S., mostly due to the difficulty of getting the correct infrastructure in place for production. And since most of the technology market is focused in the West, it is difficult for an India-based entrepreneur to form the necessary business partnerships, and find the right management, marketing and sales talent to join his or her company. (Kurian, 1997)

Within two years, Draper and Richards had reoriented the strategy for the fund to focus primarily on investments in U.S.-based firms with a large proportion of their activities in India. When the fund closed in 2000, only one of the India-based companies that had received funding was successful (Rediff Communications), whereas a majority of the U.S.-based firms in the portfolio had found liquidity. The timing was fortuitous, to be sure, but Draper attributes the successes as well to the combination of technology development in India and business, marketing, and sales operations "near the action" in the United States. Some of the firms in the fund were in the United States to start, but others were in India and Draper brought them to Silicon Valley. According to Richards, they learned about some of their best U.S.-based deals in India: "You could be sitting in Bangalore and hear about 10 great deals in Santa Clara that you could be a venture

capital investor in, but you'd hear about them first in Bangalore" (Schram, 2001, p. 66).

Draper International discovered that the ecosystem for Silicon Valley-style entrepreneurship was absent in India in the 1990s. Successful Indian and multinational software services and business-process firms had created facilities that allowed them to draw on local skill, but they remained externally oriented and isolated from the local environment. Labor mobility in regions like Bangalore was 20–30 percent in the 1990s as workers sought to maximize their earnings in a tight market, but it was not associated with the experimentation of a technical community as in Silicon Valley or Hsinchu. As a result, it simply contributed to annual wage increases on the order of 25–40 percent, as well as the increasing tendency for firms to expand to other areas (Parthasarathy, 2000).

Learning in the Indian software industry occurred primarily within the firm and through long-distance relationships, not through interactions with local customers or institutions such as universities or research laboratories—in spite of the fact that Bangalore is home to India's leading university and government research centers, including the elite Indian Institute of Science, the Center for the Development of Advanced Computing, and a range of telecommunications and defense-oriented research facilities. The public and private technology sectors, as well as industry and universities, in India remained separate spheres, with minimal professional or commercial interaction.

A study of Northern Telecom's research and development alliances with four Indian firms during the 1990s documents this pattern. These alliances were formally defined as between Nortel (as the director) and the individual firms (as software subcontractors and developers); they did not encourage interaction among the Indian partners or with local universities, and they developed products for foreign markets. Once again, these relationships were established largely through the active efforts of two senior Indian-born employees, both in the conceptualization and in the implementation. Scholars Basant and Chanadra (2004) contrast this with Nortel's collaboration with the Beijing University of Post and Telecom, which involved joint research and development based at the university and was oriented toward developing products for the domestic market. They conclude that the Indian model, while benefiting the four Indian firms, had a more limited long-term impact on the domestic skills and technological capabilities than the Chinese model.

A domestic venture capital industry emerged in India in the 1990s, but much of it was large public-sector funds or banks and multilateral institutions. Organizations like Union Trust of India Ventures or Small Industries Development Bank of India Venture Capital lacked domain knowledge or experience in software or technology-related industries. The funds, typically run by conservative financiers concerned with tangible assets and unwilling to take risks, were overwhelmingly late-stage investors in software services. In the words of A. V Sridhar, former senior manager at Wipro: "The Indian venture capitalist will not take risks in new areas, as opposed to a risk-free definitive software services market, where he is assured of quick returns and profits" (Biswas, 1998).

The supply of venture capital in India remained very small by international standards, in part because of a multiplicity of conflicting, often cumbersome and anachronistic regulations and a variety of other forms of discrimination against the industry. In 1998 only 21 companies were registered with the Indian Venture Capital Industry Association, and they had approximately $700 million available for investment. This compared to Israel's 100 funds with $4 billion available for investment in 1999 and Taiwan's 110 funds with $1.32 billion investments.

A few domestic technology start-ups survived the environment of the 1990s, as well, but none experienced the successes of start-ups in Taiwan or Israel at the time, although most were started by returnees from the United States (Desai, 2003). For example, Silicon Automation Systems (SAS) was started in Silicon Valley in 1989 by four U. S.-educated Indian engineers. They bootstrapped the venture, and in 1991 they moved part of the operation to Bangalore while seeking to maintain a strong research and software and hardware design (as opposed to software services) orientation. In 1997 SAS was a small, privately held firm with about three hundred employees and $10 million in exports. Another returnee, Pradeep Singh, started the software services firm Aditi in Bangalore in 1994, after nine years at Microsoft. With headquarters in Seattle, Aditi relied on Microsoft as a stable client while also seeking to grow the business. During the 1990s most of the other returnee start-ups in Bangalore and elsewhere in India—a small number, to be sure—either failed or stagnated.

Growth and Constraint

The boom in the U.S. technology sector in the late 1990s had contradictory consequences for India. On one hand, labor shortages resulted in an

increase in the quota for temporary visas granted on the basis of skill, with 124,697 Indian nationals gaining approvals of H1-B status in 2000 alone, representing nearly half (48 percent) of all visa approvals. The next largest sending country was China, accounting for only 9 percent of H1-B recipients (22,570). Desai (2003) estimates that at least a half-million Indian programmers received U.S. visas (of all sorts) between 1996 and 2001.

Labor shortages also contributed to the growing numbers of cross-regional start-ups between Silicon Valley and India. Rakesh Mathur, who worked for Intel in Silicon Valley for many years before starting three successive successful technology companies—Armedia, Junglee, and Stratify, explains: "The key constraint to starting a business in Silicon Valley in the late 1990s was the shortage of software developers. I realized that I could go to India. All three of my start-ups had design centers in Bangalore but were registered as American technology companies" (Mathur, 2002). Mathur made six trips to India in 1999, and in 2000 his firm Stratify established a hundred-person operation in Bangalore.

Mahesh Veerina, who started self-financed Internet technology firm Ramp Networks in 1993 with two friends, reports that they were quickly running out of money and couldn't afford to pay for local engineers, so they hired programmers in India for one-quarter of the Silicon Valley rate. The firm's engineers reported that they were able to cut development time in half because Indian team worked while the U.S. team slept. By 1998 the firm had sixty-five employees in Santa Clara and twenty-five in India—and required that every engineer spend at least a couple of weeks working in the other country (Thurm, 1998). While data are not available, it appears that this cross-regional business model was increasingly common in the late 1990s.

While firms like Ramp Networks and Stratify represent the standard model for building a cross-regional company, starting in the United States and tapping talent in India, Draper International pioneered the reverse strategy. In 1997 the venture capital firm identified and recruited A. V Sridhar, a senior manager at Wipro who had never worked outside India, to Silicon Valley to start a company in the data mining field. Sridhar quickly identified a marketing team, including Sanjay Anandaram, who was already working for Wipro in Silicon Valley, and senior managers with experience at Oracle India and IBM Research. The start-up, Neta Inc., developed Internet personalization software and was acquired two years later by Internet portal Infoseek (now Go.com). Sridhar explains why he moved to Silicon Valley to start the business: "To create a successful

company, one has to be real close to the market. One has to be in a place which supports the creation of new technologies as a daily affair" (Biswas, 1998).

Permanent returnees to India from Silicon Valley remained few and far between throughout the 1990s, but the professional and personal networks linking Indians in Silicon Valley to family members, friends, and colleagues at home combined with access to e-mail and low-cost travel and phones to generate an unprecedented rate of information exchange between the United States and India. In the words of a Silicon Valley engineer:

> I go back to India two or three times a year because of my work and there are parts of India where you take a train and go over there and they don't even have a rickshaw or a cab to take you to your destination. You have to walk. But everybody in the small town knows exactly what the job situation is in Silicon Valley. They know the H1-B quota level, when it is filled, when it is open again. They know exactly what kinds of skills are required in Silicon Valley, not even in California, just Silicon Valley. (Santa Clara County Office of Human Relations, 2000)

Web sites like www.return2india.com, Non-Resident Indians Online (www.nriol.com), and www.siliconindia.com became increasingly popular among U.S.-educated Indians. They traveled home frequently, some returning home to get married (often following the traditional practice of arranged marriage), and seemed torn between the familial and cultural pull of home and what they regarded as superior professional and economic opportunities in the United States.

The potential for substantial returns has attracted a new generation of venture capitalists. A group of new $50 million to $100 million funds, typically from established markets such as Walden International and E-Ventures (Softbank), targeted early-stage investments in India. Corporate venture capital also became active in Indian technology regions. Intel Capital committed to invest $100 million in Indian IT start-ups during 2000, and Computer Associates began investing actively as well. Traditional sectors that had earlier shunned the software industry became interested in investing because of the high valuations. Fund managers began to tap India's high-net-worth individuals and family firms for capital. Total venture capital investments in India reached $1160 million in

2000. The level declined to $937 million in 2001 and dropped off further in 2002 and 2003.

A new generation of cross-border investors with accumulated experience in both India and Silicon Valley has emerged, as well. The JumpStartUp Venture Fund was established in Bangalore in July 2000 by three veterans of the IT industry: Kiran Nadkarni had fourteen years of experience with venture capital in India, serving most recently as the Draper International partner in Bangalore; K. Ganapathy Subramanian came from ICICI Venture, a leading venture capital firm in India; and Sanjay Anandaram had worked at Wipro in both Bangalore and Silicon Valley and then at start-up Neta Inc. The $45 million fund was targeted at early-stage information technology start-ups and had funding from both institutional investors (including Silicon Valley Bank) and individuals, including Bill Draper and successful Indian executives in the United States (contributing 20 percent of the fund).

In 2002 JumpStartUp moved its headquarters from Bangalore to Santa Clara, California, in order to shift its investment strategy from an India-focused fund toward "U.S.-India cross-border investments." The partners realized that their small fund was not sufficient to support early-stage start-ups from the ground up in the environment of the early 2000s, when outside investors were reluctant to contribute to cash-poor companies. The new strategy recognizes that Silicon Valley's cash-strapped new ventures must increasingly set up engineering centers in India. JumpStartUp envisions a role as co-investor with established venture capital firms in order to help portfolio companies set up engineering teams as well as design, deployment, and support functions in India. In the words of Nadkarni: "It is very hard for companies started by non-Indians to think of working out of India, unless they have done it in the past. Unless the founders have intent to do things out of India …. When a startup decides to establish a development center in India, invariably you will see that one of the founders is an Indian" (Shankar & Sundaram, 2003).

Most venture capital in India remains focused on later-stage investments in software services firms, but venture capital firms with Indian fund managers who bring technology investment and entrepreneurial expertise have targeted the early-stage U.S.-India connection as well, including WestBridge Capital Partners and Artiman Ventures. The economic logic for this structure appears compelling: a company that would need

$10 million to $15 million in its first round of funding in Silicon Valley might hire a comparable engineering workforce in India for only $2 million to $3 million.

5 Conclusion

The old pattern of one-way flows of technology and capital from the core to the periphery is being replaced by a far more complex and decentralized two-way flow of skill, capital, and technology between differently specialized regional economies. Silicon Valley is now at the core of this rapidly diversifying network of economies because it is the largest and most sophisticated market as well as a leading source of new technologies. However, this too could change: the relationships between these emerging technology regions are multiplying and new markets are opening up in China and India that promise to further transform the dynamics of the world economy.

Regions like Hsinchu, Bangalore, and Shanghai are not replicas of Silicon Valley—although institutions and professional service providers from that region are fast expanding into these new environments. These new regional economies have co-evolved with the Silicon Valley economy. Firms in these regions do not typically seek to compete directly with Silicon Valley producers. They focus instead on developing capabilities in areas that U.S. producers are not pursuing; and over time they have transformed activities once regarded as mundane and low-tech into more efficient and dynamic sectors. So while producers in these regions entered the global market by providing low-cost skill, each has developed specializations that add distinctive new value to electronic products and systems. Taiwan was known in the 1980s for its cheap PC clones and components; today it is recognized for the flexibility and efficiency of its IC and electronic systems producers. China was known in the 1990s for me-too Internet ventures; today Chinese producers are poised to play a lead role in developing wireless technology. In the 1990s India was a provider of labor-intensive software coding and maintenance; today local companies are managing large-scale software services projects for leading global corporations. Israel was a low-cost location for research in the 1980s; since then, local entrepreneurs have pioneered sophisticated Internet and security technologies.

These developments explain why Silicon Valley-based firms are active participants in all of these regions as investors and as partners, not simply as competitors. An established firm, like Cisco, designs and sources critical

parts of its operating system software from India and application-specific integrated circuits from Israel for its high-end routers, and the manufacturing of most of its hardware from Taiwan and China. It also invests in start-ups with promising technologies in these locations. On the other hand, a start-up (or "micro-multinational") like July Systems obtained venture capital from firms based in the United States, Taiwan and China, and India, and its products will likely incorporate components from all these locations as well as being targeted at all their markets.

The new technology regions have all become high-wage, high-cost locations in their national economies, yet they continue to spawn start-ups and attract new firms, and most existing producers continue expanding locally. This growth is reminiscent of the continued clustering of computer and communications-related firms in Silicon Valley during the 1980s and 1990s; these areas now boast a regional advantage that compensates for their high costs. Silicon Valley producers no longer view locating or sourcing from India or China as an efficient way to reduce costs; rather, they argue that the only reason to work with producers in those locations is to gain access to the talent.

U.S. technology producers benefit directly from the development of these specialized technology regions. They now look to their counterparts in Taiwan and China, India, and Israel not simply for low-level implementation but increasingly for co-development and co-architecting of products and components. In addition, firms in the new technology regions are increasingly partnering with one another as well as with firms from Silicon Valley, as when a Taiwanese semiconductor firm invests in Israeli start-ups specialized in digital speech processing chips, or when an Israeli company contributes intellectual property components to a chip design firm based in India. These collaborations deepen the capabilities of each of the partners and over time can support a process of reciprocal innovation and upgrading in the respective regions.

The nature of the firm is being redefined through this process as start-ups and investors as well as established enterprises move across geographies and open their boundaries in new and different ways. The corporate hierarchies and vertical integration of the twentieth-century mass production corporation are being replaced by horizontal networked organizations that flexibly recombine resources by drawing on the specialized and complementary capabilities of producers located in distant regions. This often involves extensive information exchange and collaboration between customers and suppliers who are located at great distances. This suggests

that the region and its distinctive capabilities and resources, rather than the firm, are increasingly defining the contours of the global production.

A Model for Others?

This is not to suggest that all developing economies are positioned to reap the benefits of brain circulation and peripheral entrepreneurship. This opportunity is benefiting countries that have invested heavily in higher education, typically technical education, and are politically and economically stable enough that immigrants will consider returning home. Some of the largest technically skilled immigrant groups in Silicon Valley have not built business or professional connections to their home countries for political reasons. Most of the region's Iranian and Vietnamese immigrants, for example, are political refugees and hence not inclined to return to countries that, in any case, lack the economic stability needed for technology investment or entrepreneurship. This criterion applies in varying degrees to many of the developing economies that have technically skilled communities in the United States and at home, including Russia, parts of Eastern Europe, and Latin America. It is possible that urban areas like Saint Petersburg or Buenos Aires will become more attractive to returning entrepreneurs in the future as their economies develop and eventually provide greater professional opportunities for returnees. However, large parts of Africa and Latin America lack the skill base or political openness to become attractive environments for technology entrepreneurship.

Industrialized economies in Europe and Asia face different challenges. While they typically boast well-developed skill bases, research capabilities, and sophisticated infrastructure, technical skill is significantly more costly in these economies, particularly relative to recently developed economies like Taiwan and Israel. The financial and corporate structures in industrialized economies often discourage technology entrepreneurship, as well. In Japan and France, for example, the privileged relationship of large corporations, banks, and the state limit the opportunities for outsiders (returning entrepreneurs) to raise capital or develop domestic markets. Continuing discussions in Europe of creating industrial champions suggests that policy there may continue to reinforce, rather than dismantle, these relationships.

The structure of labor markets mirrors this configuration: the technical elite in countries like Korea and France move automatically into high-status positions at the top of the large corporations and banks or the civil

service. They have little incentive to study or work abroad, and face significant opportunity costs if they do. This disincentive helps explain why relatively few pursue graduate education in the United States, and why those who do frequently return home directly upon graduation. Those who end up in Silicon Valley are less likely to be the top of their class, and they gain little respect, let alone access to capital, by returning home.

In many Asian countries, government support for large-scale, capital-intensive investments in the 1970s and 1980s, either by domestic corporations (Korea) or by multinationals (Singapore), have created inhospitable environments for entrepreneurial experimentation. One indication of this is data on the sources of innovation. South Korea's *chaebol*, or large business groups, accounted for 81 percent of all U.S. patents earned in Korea in the 1990s compared to only 3.5 percent earned by business groups in Taiwan (Trajtenberg, 2000). Likewise, in South Korea the top fifty assignees accounted for 85 percent of all U.S. patents, with Samsung alone accounting for 30 percent, while Taiwan's top fifty assignees accounted for only 26 percent of all U.S. patents. This decentralization of innovative capabilities was reflected in a substantially higher rate of patenting in the late 1990s, with Taiwan earning 17.7 patents per US$ billion exports compared to 11.6 in Korea (Mahmood & Singh, 2003).

Evidence from the Korean and Japanese communities in Silicon Valley suggests that even when entrepreneurs are successful in the United States, they often lack opportunities for alliances and partnerships at home. It is very difficult for these small firms to collaborate on equal terms with the giant *chaebol* and *keiretsu* (industrial groupings in Japan). This contrasts to the Taiwanese technology entrepreneurs in Silicon Valley whose firms have grown up through partnerships with their counterparts, often also entrepreneurs, at home. And as with European cases, Japanese and Korean graduate students earning degrees in the United States are typically lured back home immediately by attractive opportunities in government, large corporations, or universities.

Another group of developing economies has grown since the 1970s as recipients of manufacturing investments by United States, Japanese, and European technology corporations. These investments, which were targeted at low-wage locations including Singapore, Malaysia, Scotland, and Ireland, have contributed to the development of the supplier infrastructures and skill base needed to master high-volume manufacturing of electronic components (McKendrick et al., 2000; Rasiah, 2001). They have also contributed to substantial improvements in standards of living.

However, the leading recipients of foreign direct investment remain technological laggards. The rate of patenting, normalized by either population (Trajtenberg, 2000) or exports (Mahmood & Singh, 2003) in Singapore, Malaysia, Hong Kong, and Ireland since 1970 remains a small fraction of that observed in Taiwan and Israel. Seven of the top ten patent recipients in Singapore, for example, were foreign multinationals or organizations, accounting for 46 percent of all U.S. patents between 1970 and 2000 (Mahmood & Singh, 2003).

Recent policy changes, such public support of venture capital, have not been sufficient to transform domestic institutions and capital and labor markets. In these nations, skilled workers prefer stable, corporate employment, and start-ups lack access to financial and technical resources as well as markets. The 2000 *Global Entrepreneurship Monitor*, for example, found that in spite of higher than average GDP growth, Singapore had "one of the lowest rates of entrepreneurial activity" of more than twenty countries it studied (Reynolds et al., 2001). Returning engineers to these nations have not made a significant impact on their home countries.

The creation of a transnational community is a two-way process. While policymakers and planners can encourage cross-regional connections, they cannot create or substitute for transnational entrepreneurs and their decentralized networks. Government agencies from Singapore, Hong Kong, and Japan regularly sponsor networking events for their expatriates in the Bay Area as a way to recruit return entrepreneurs and investments. However, the absence of entrepreneurial collaborators at home means that these agencies can provide incentives and information but not access to partners with an interest in jointly transforming the home environment. Governments cannot by themselves insure the preconditions for return entrepreneurship; this is inherently a process of collaborative institution building that takes both local knowledge and understanding of global technology markets and networks.

Cross-regional networks develop only when skilled immigrants are both willing and able to return to their home countries for business in large enough numbers to create close links to the technical community in the home country. The receptiveness of the home country depends upon factors such as political stability, economic openness, and level of economic development. It often builds on multinational investments in research and development that have contributed to a developing local skill base and infrastructure that supports entrepreneurship. The critical variable is the possession of political leaders willing to collaborate with

returning entrepreneurs to develop a shared vision and remove institutional and political obstacles to entrepreneurship-led technology growth.

The Future of China and India

India and China, both very large developing countries that opened their markets to the global economy in the last two decades, have less-defined trajectories. The political and economic environment in these countries is complex, multilayered, regionally differentiated, and continually in flux. Both countries have attracted substantial foreign investments based on the combination of large technically skilled workforces and low wages. For both, brain circulation with an overseas community in Silicon Valley has been an important factor in attracting foreign investment. Policymakers from both countries have established ties to their large technical communities in Silicon Valley. Local and regional governments in both countries compete intensely for technology activity, and in each technology investment and entrepreneurship remain concentrated in a handful of urban centers.

The leading Indian technology firms today specialize in exporting software services to American and European corporations and are concentrated in a handful of large urban centers in southern India. The leading Chinese information technology businesses focus primarily on manufacturing for the domestic market, and are concentrated in the leading urban areas along the east coast of China—where the export products are overwhelmingly produced by foreign companies, mainly from Taiwan. China has benefited as well from growing ties to Taiwanese businesses, in spite of the political tensions across the Straits. Both cases show signs of the dynamic growth of cross-regional start-ups that will become competitive in global markets or grow sizable new markets. These firms are primarily started by returnees from Silicon Valley, and they have the potential to provide the basis for a decentralized entrepreneurial industrial system. However, it is still very early in the process: neither country has seen a full cycle of entrepreneurial success and reinvestment.

Some differences can be traced to institutional factors. In China, governments at all levels have channeled substantial investment into developing an urban infrastructure and growing the domestic market. The selection of information technology industries as a national priority has also triggered vastly increased investments for research and development, technical education, the adoption of information technology in both

public and private sectors, and the construction of a seemingly unlimited number of science parks. Political control by the Communist Party, ongoing corruption, and chronic weaknesses in financial markets continue to create uncertainties about China's future.

India, by contrast, made expansion of export earnings a top priority in order to improve its balance of payments at a time when distrust of bureaucratic intervention (based primarily on the failures of the postwar "license Raj") led to the gradual embrace of liberalization and a largely hands-off approach to industrial development. Nonetheless, bureaucratic obstacles remain significant and have limited efforts to improve the domestic infrastructure—from roads and airports to energy supply and telecommunications. The risk of enclave-like development is quite real in India, where thriving technology districts like Bangalore appear better connected to Silicon Valley than to their impoverished domestic hinterlands. Unlike in China, where the government has used its resources actively to insure that investment and growth continue in the west as well as the wealthy eastern regions, the distrust of the public sector in India constrains the educational and infrastructural investments needed to spread the benefits of growth.

Both India and China are increasingly important actors in the global economy and particularly in technology industries. There are, however, many unanswered questions about their developmental trajectories. The technology sectors remain a small proportion of overall economic activity in both countries, and remain highly geographically concentrated. However, either could have a disproportionate impact over the long term. We can only speculate here based on the experiences of other countries and the initial experience of each over the past two decades.

Technology markets are shifting quickly, however, with demand growing rapidly outside the United States. While the Asia-Pacific region (excluding Japan) accounted for only 6 percent of semiconductor consumption in 1985, in 2000 it accounted for 25 percent, and is predicted to grow to 46 percent by 2010 (with China accounting for 16 percent). China and India are now the fastest-growing markets for wireless technologies. In 2001, 400 million mobile handsets were sold in China. This is creating new opportunities for producers in these economies to develop products for their domestic markets. These markets are also likely to provide product and design inputs for the United States. SMS messaging on cell phones, for example, originated in Asia and Europe but has become increasingly popular with American youth. Likewise, the fuel cells

developed for low-cost, long-lasting energy solutions for China and India could also change the economics of similar products in the United States.

Today the West and Japan produce more than half of global output but account for less than 15 percent of the world's population. This will change decisively in the next fifty years; some observers predict that customer demand from India and China will dominate global markets within a decade, and that the Chinese and Indian economies will be bigger than the U.S. economy by 2050 (Wilson & Purushothaman, 2003). New regional economies will certainly develop sufficient technology capabilities to participate in global networks, as well. They will likely share with their predecessors a history of long-term investments in education and research, as well as the institutional openness that derives from the absence of organized interests from an earlier industrial era. Silicon Valley's role as the dominant technology center will continue to diminish. This does not mean that the region will decline, only that it will become one of many nodes in an open and distributed global network of differently specialized and complementary regional economies.

References

Basant, R., & Chanadra, P. (2004). Capability Building and Inter-Organizational Linkages in the Indian IT Industry. In A. D'Costa & E. Sridharan (Eds.), *India in the Global Software Industry.* Macmillan India.

Biswas, S. (1998). Leading the VC Wave in India. *Siliconindia*, May/June. Retrieved September 19, 2004, from http://web.archive.org/web/19990218061723/www.siliconindia.com/MayJune98wave.html

Chan, W.-C., Martin, H., & Stephen, S. (2002). The Greater China High Tech Highway. *The McKinsey Quarterly*, 4.

Chwang, R. (2003, May 29). *Venture Investment in Greater China—Opportunities and Challenges.* Presentation at Stanford University.

Clendenin, M. (2003). Befriending the Dragon. *Electronic Engineering Times*, October 31. Retrieved September 2, 2004, from http://www.eetimes.com/futureofsemis/

Desai, A. V. (2003). The Dynamics of the Indian Information Technology Industry. Center for New and Emerging Markets, London Business School, Apr. Retrieved September 6, 2004, from http://www.london.edu/cnem/Faculty/S_Commander/india27603.pdf

Hobday, M. (1995). East Asian Latecomer Firms: Learning the Technology of Electronics. *World Development, 23*(7), 1171–1193.

Kurian, T. (1997). In the Footsteps of Giants. *Siliconindia*, November. Retrieved September 19, 2004, from http://web.archive.org/web/19990508224300/www.siliconindia.com/3rdGiants.html

Mahmood, I. P., & Singh, J. (2003). Technological Dynamism in Asia. *Research Policy, 32*, 1031–1054.

Mathur, R. (2002). *Presentation at Forum on Local And Global Networks of Immigrant Professionals*. Public Policy Institute of California.

McKendrick, D. G., Doner, R. F., & Haggard, S. (2000). *From Silicon Valley to Singapore: Location and Competitive Advantage in the Hard Disk Drive Industry*. Stanford University Press.

Parthasarathy, B. (2000). Globalization and Agglomeration in Newly Industrializing Countries: The State and the Information Technology Industry in Bangalore, India. Ph.D. dissertation, Department of City and Regional Planning, University of California, Berkeley, Spring.

Rapaport, H. (2002). *Who Is Afraid of the Brain Drain? Human Capital Flight and Growth in Developing Countries*. Stanford Institute for Economic Policy Research, Policy Brief. April.

Rasiah, R. (2001). Politics, Institutions and Flexibility: Microelectronics Transnationals and Machine Tool Linkages in Malaysia. In F. C. Deyo, R. F. Doner, & E. Flershberg (Eds.), *Economic Governance and the Challenge of Flexibility in East Asia*. Lanham, MD.

Reynolds, P. D., et al. (2001). *Global Entrepreneurship Monitor, 2000 Executive Report*. Babson College.

Santa Clara County Office of Human Relations. (2000). *Building Borders in Silicon Valley: Summit on Immigrant Needs and Contributions*. Santa Clara County Office of Human Relations.

Saxenian, A. (2003). Government and *Guanxi:* The Chinese Software Industry in Transition. Paper Presented at the conference "Global Software from Emerging Markets: An Engine for Growth?" Centre for New and Emerging Markets, London Business School, May 12.

Schram, A. (2001). Venture Guru. *Siliconindia*, September: 66–68. Retrieved September 7, 2004, from http://www.siliconindia.com/Magazine/fullstory.asp?aid=LCX354848695

Shankar, P., & Sundaram, K. (2003). The Cross-Border Nadkarni. *Siliconindia*, February 2. Retrieved September 23, 2004, from http://www.siliconindia.com/person_ftweek/pow.asp?id=133

Thurm, S. (1998). High-Tech Firms Big and Small Go Global. *San Jose Mercury News*, May 9.

Todaro, M. P. (1985). *Economic Development in the Third World*. Longman.

Trajtenberg, M. (2000). R&D Policy in Israel: An Overview and Reassessment. BER Working Paper No. w7930, National Bureau of Economic Research, Cambridge, MA. October 2000. Retrieved September 14, 2004, from http://papers.nber.org/papers/w7930.pdf

White, S., Gao, J., & Zhang, W. (2002). China's Venture Capital Industry: Institutional Trajectories and Systems Structure. Paper Prepared for INTECH, International Conference on Financial Systems, Corporate Investment in Innovation and Venture Capital, Brussels, November 7–8.

Wilson, D., & Purushothaman, R. (2003). "Dreaming with BRICS: The Path to 2050" Goldman Sachs Global Economics Paper No. 99. Retrieved May 31, 2005, from www.gs.com/insight/research/reports/99.pdf

Woetzel, J. R., & Chen, A. C. (2002). Chinese Chips. *McKinsey Quarterly*, *2*. Retrieved September 19, 2004, from http://www.mckinseyquarterly.com/article

Xiao, W. (2002, September 30). The New Economy and Venture Capital in China. *Perspectives, 3*(6). Retrieved September 7, 2004, from http://www.oycf.org/Perspectives/18_093002/Economy_Venture_China.htm

International Competition Intensified: Job Satisfaction Sacrificed?

Barbara Dluhosch and Daniel Horgos

1 Introduction

Throughout the last decades, many countries have witnessed a substantial increase in outsourcing-offshoring.[1] Production became virtually global with many steps and tasks in the production process having been moved abroad. According to the OECD-WTO October 2015 edition on trade in value added (TiVA), which provides the most recent numbers of the OECD-WTO project on measuring global value chains (GVCs), foreign value added in exports climbed in the US from 11.46 (1995) to 15.03%

[1] We follow the convention established since (UNCTAD, 2004, p.148) that defines offshoring as imports of intermediates, regardless of whether they are produced by a foreign affiliate or a third party but in any case produced abroad whereas outsourcing is intermediate production of domestic or foreign origin.

B. Dluhosch • D. Horgos (✉)
Department of Economics, Helmut Schmidt University – University of the Federal Armed Forces Hamburg, Hamburg, Germany
e-mail: international.econ@hsu-hh.de; daniel@horgos.de

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_7

(2011).[2] Over the same period, TiVA rose in Germany from 14.86 to 25.54%, in France from 17.29 to 25.13%, in Spain from 19.16 to 26.88%, and in the UK from 18.25 to 23.05%. While the US and Australia are at the lower end of the GVC spectrum, Eastern European countries, such as the Czech Republic, Hungary, and the Slovak Republic, are at the upper end, with a fraction of foreign content in exports of more than 45%. All trade within GVCs accounted for almost half of world trade in goods and services in 2011 (WTO, 2015). Although a number of indicators lately show a slowdown in GVC-related trade for some countries, the share of intermediates in world trade continues to be high, especially with respect to trade in services, which grew well above average (UNCTAD, 2015, pp. 4, 22 et seq., and Tables in Chap.5).[3]

The rise in offshoring sparked an intense debate as to possible labor market implications. While not undisputed, many studies find that, in the advanced economies, the process of offshoring has contributed to a relative decline in the demand for low-skilled labor as low-skill intensive parts in the production were substituted by cheaper imports from abroad. Almost all of the studies, however, exclusively concentrated on employment and income effects of this process. In particular wage inequality across workers of different skill levels, occupations, and industries were the focus of interest (see, e.g., Becker & Muendler, 2015 and Ebenstein et al., 2015 on country-specific perspectives or Hummels et al., 2016a for a survey).

Surprisingly less research has been done on the effect of offshoring on subjective well-being (SWB). There is a lack of research on the globalization of production and SWB, despite the fact that recent research has shown that income, though important, is just one aspect of life among several that feed into an individual's SWB. Although there is some controversy surrounding the exact channels and impacts, findings suggest that income need not even be positively correlated with SWB (e.g., Easterlin,

[2] See http://stats.oecd.org/Index.aspx?DataSetCode=TIVA2015_C2 (as of March 30, 2017) for details.

[3] With measurement problems abound, indicators differ as to the extent and the dynamics of offshoring and vertical specialization across international borders. Numbers thus have to be taken with a grain of salt, see, for instance, Sturgeon (2013) and the survey by Amador and Cabral (2016) on measurement issues. While high with approximately 32%, China fails to meet the trend upwards. For a possible explanation drawing on the importance of the domestic market and economic policies, see Brandt and Thun (2016). There is no dispute though about the extent being substantial.

1974; Veenhoven & Hagerty, 2006; Epstein, 2008; Helliwell & Barrington-Leigh, 2010; Diener et al., 2013). However, SWB in itself is to be considered an amalgam of different perceptions (Hsieh, 2016). When thinking about how offshoring might affect SWB, job satisfaction is certainly an issue that immediately comes to one's mind. Naturally, labor economists have for a long time identified job satisfaction as important matter (e.g., Hamermesh, 2001; Sousa-Poza & Sousa-Poza, 2003; Lévy-Garboua & Montmarquette, 2004; Clark et al., 2009; Card et al., 2012). Although, apparently, job satisfaction is an important issue if judged by the sheer number of these contributions to labor market research, the nexus between international trade and job satisfaction has remained largely unexplored territory.

In this chapter, we try to fill the gap by illuminating the impact of trade and offshoring on job satisfaction. To this end, we draw on a large sample of data from the International Social Survey Program (ISSP), the Penn World Tables (PWT), the World Development Indicators (WDI), trade freedom data by the Heritage Foundation, the Fraser Institute, and the KOF globalization index taking in some 31 countries and 24,619 individuals. The empirical analysis focuses on the effects on job satisfaction as a subjective variable. We also rely on subjective information when capturing the characteristics of the workplace.[4] In order to understand the process and how it is embedded into income effects, we sketch out a small model that is capable of explaining consequences of trade and offshoring for job satisfaction in their various dimensions, in particular real income effects and job characteristics.

Our results run counter to what might be intuitively considered a negative relationship in the sense that increased competition from abroad and the rise in offshoring might be thought of lowering job satisfaction. Instead, the picture seems to be much more diverse with some trade indicators being associated with a decline in job satisfaction while others report an increase. However, most notably, in countries with more intermediate products being produced abroad (thus resulting in a high offshoring index), job satisfaction scores significantly higher than elsewhere. Our

[4] The OECD (see Cazes et al. 2015 for an introduction) also offers data on (subjective) job quality, however, with a large overlapping as to countries. For reasons of consistency and because the ISSP offers a wider set of countries with more variance in the data, we opt for the ISSP. In fact, the OECD job quality data (see http://stats.oecd.org/Index.aspx?DataSetCode=JOBQ, accessed March 30, 2017) are based to a sizable extent on the ISSP. As to the associated measurement issues, see Muñoz de Bustillo et al. (2011).

findings prove to be robust in several cross-checks with respect to the data as well as the econometric setup. Digging deeper into the reasons for the positive nexus by means of a further disaggregated approach, a socio-psychological dimension appears to be important next to the pure economic effects on wages, income, and consumption. Beyond improvements of more general social and psychological aspects like trust, self-categorization, knowledge, or the opportunity for better career development, all of which come along in tandem with offshoring activities, we are able to trace the effect back to an alteration of job characteristics in the process of offshoring and the associated technological change. If, for example, monotonous, repetitive tasks are offshored or disappearing altogether because of associated changes in production, remaining tasks may be more inspiring, attractive, and interesting. If so, job satisfaction may score higher. If the two components, pecuniary and non-pecuniary aspects of job satisfaction, enter with different signs, socio-psychological aspects of the workplace may compensate for negative income effects.

Previous contributions that explored how the process of internationalization is individually perceived concentrated on individual preferences over international trade and immigration policies (see, e.g., Scheve & Slaughter, 2001; Hanson et al., 2007; Khun et al., 2015); others analyzed the effects of globalization on individual well-being, which seem to exert an SWB effect separate from individual income and distributional consequences. According to Bjørnskov et al. (2008), openness to trade, defined as exports plus imports over GDP, is among the small number of variables that robustly affect individual life satisfaction. Also, trade policy, as measured by various trade freedom indices, is found to significantly affect individual well-being if non-tariff barriers are considered next to the average tariff rate (Dluhosch and Horgos, 2013). In contrast to actual trade flows, this channel can be interpreted as the option value of trade. Turning to more specific aspects of globalization, Scheve and Slaughter (2004) found that FDI is a key variable significantly increasing worker insecurity. We account for these channels by tracking these various facets of globalization (openness, trade freedom, outward FDI) in addition to offshoring proper, thus moving beyond those contributions.

As to aspects of the workplace in relation to the globalization of production, Geishecker (2008) lends support to the notion that international outsourcing lowers individual employment security significantly. The effect varies with job duration but not with the skill level of the employee. Job loss fears though seem to depend on the region to which production

is being offshored: outsourcing to high-wage countries significantly decreases job loss fears while outsourcing to low-wage countries significantly increases those fears (Geishecker et al., 2012). Geishecker et al. (2012) also presents empirical evidence of job security being strongly correlated with individual well-being. Colantone et al. (2015), by drawing on the British Household Panel Survey, present indication for the UK that competition and the associated adjustment costs cause mental stress. Hummels et al. (2016b) even find adverse health effects of exports in Danish matched worker-firm data.

In fact, as to previous research, the majority has focused on the issue of job (loss) fear and issues of job security. This is certainly an important aspect that warrants discussion, and, despite the number of contributions, it may still provide fertile ground for research. For instance, one may ask how perceptions deviate from reality as to job security, that is whether there is a cognitive dissonance on these matters or not (e.g., perceptions may underrate or overestimate the risk of job loss). On a similar account, perceptions may be driven by a "survivorship" bias, to the effect that they do not reflect the true risks.[5] Other than previous contributions, this would draw, in addition to the primary survey data, on comparisons of different data sets in order to detect inconsistencies. And, other than previous contributions, it would thus address explicitly the details of the psychological-social dimension involved. And, it probably would deliver on how those misperceptions can be addressed either at the firm level by the management or even at the policy level by appropriate social security design.[6]

This chapter, by contrast, is more directly related to issues concerning how the quality of the workplace and job characteristics are perceived and thus with overall job satisfaction rather than the particular issue of job insecurity, the fear of losing one's job, or the mental stress associated therewith. Consistent with the focus on how workplace characteristics are perceived, we follow a methodological-individualism approach in that we

[5] We owe these points to a discussion with one of the anonymous referees.

[6] The dissonance may run both ways. While intuition (and, e.g., research by Kahneman and Tversky 1979) suggests that risks may be overestimated by risk (loss) aversion as is commonly presumed, Gallup surveys (Newport 2012) suggest the opposite as to job loss fears. From the perspective of cognitive dissonance, the latter may be explicable by keeping the psychological cost low(er) in case of eventual job loss. Similarly, a survivor syndrome may be involved as workers reduce identification and effort when facing the risk of losing their job (e.g., Datta et al., 2010).

exclusively draw on data on self-reported well-being. These dimensions of job satisfaction, although an obvious candidate when it comes to subjectively evaluating the trend toward the globalization of production, have thus far been largely left out. Böckerman and Maliranta (2013) is one of the few exceptions. Exploring matched employer-employee data for Finland, they see a "silver-lining" in that outsourcing has some positive effects on employee well-being via intrafirm occupational restructuring. Because their results are based on those remaining employed and the case of Finland in particular, they are reluctant, however, to draw wider conclusions as to possible economy-wide effects or whether effects reach beyond the case of Finland. Also, they do not differentiate between outsourcing within and between countries, whereas our focus is specifically on the impacts of international outsourcing. We also widen the perspective by use of ISSP data on more countries and thus a more comprehensive sample of individuals. Still, the positive nexus between (international) outsourcing and job satisfaction holds.

Exploring in more detail the question as to why there is obviously also a dimension of job satisfaction that is (subjectively) positively evaluated with international outsourcing reveals that there are some job characteristics that change as tasks are being offshored. Those that stay employed perceive them as improving: remaining employees perceive their job to be more interesting, exhaustive, but less stressful. Notably, results hold when accounting inter alia separately for how issues of job security affect SWB. The offshoring-related changes in job characteristics can even dominate income effects. In this respect, our findings are in line with Yeh (2015), who, based on data from the 2005 ISSP (Work Orientations III), provides empirical evidence for East Asia of job quality being more important than earnings for job satisfaction. Speculating on the reasons, she considers this specifically tied though to the guanxi tradition in East Asia which is absent in more Western societies.

With respect to the theoretical explanation, we will borrow from Dluhosch and Hens (2016). However, our model is enriched with a socio-psychological dimension capturing non-pecuniary aspects of the workplace, like trust, knowledge, group membership and identification, career development, and job characteristics in the more narrow sense of the term. Offshoring activities change these job characteristics. Consequently, our model enables us to examine offshoring-related changes in job satisfaction. With respect to the way these non-pecuniary aspects are incorporated in the model, we follow Lü et al. (2012), who consider well-being a

weight function of objective and subjective inputs. However, while their focus is on non-pecuniary aspects related to income comparisons and how income inequality is perceived (see also Clark and Oswald, 1998 for various specifications as to relative positions), ours is on the fragmentation of the production process and on offshoring-related changes in workplace characteristics. It thus offers a novel explanation as to how offshoring impacts job satisfaction besides any consumption effects it might have on subjective well-being.

The remainder of the chapter is in three major parts. In Sect. 2, we present results from our theoretical approach to explaining how offshoring affects job satisfaction via several channels. A small simulation study gives a first impression on net effects of pecuniary and non-pecuniary aspects of offshoring and how the separate effects relate to each other by parameterizing their composite. At the end of Sect. 2 we formulate two testable hypotheses to guide the empirical section. Section 3 then is devoted to the empirical analysis. After introducing the data, we employ a cross-section logistic regression model to estimate the impact of various dimensions of trade (and offshoring in particular) on job satisfaction. In Sect. 4, we perform multiple robustness checks on the empirical setup. Section 5 summarizes and finishes off with some thoughts on future research.

2 Offshoring and Individual Job Satisfaction: A Theoretical Framework

In order to illuminate and better understand the channels via which offshoring might affect job satisfaction in its two dimensions, (semi-)pecuniary (i.e., pleasure from consumption possibilities) and non-pecuniary (i.e., subjective well-being at the workplace), we need to track down what drives offshoring and how this relates to the two dimensions. Exploring the details shows that via a macro-micro link, the underlying processes with respect to the workplace are inevitably intertwined with income effects (and the individual reactions to them), which in turn affect (the extent of) offshoring (which feeds back on the workplace). To fully capture how they relate to each other requires examining the drivers and the consequences of offshoring in some detail. Analyzing the drivers and channels also serves to set up the empirical model properly and identify potentially important variables. In what follows, we will concentrate on

the understanding of the nexus rather than the theoretical details of a fully developed model.

Suppose the economy is populated by two kinds of individuals, low- and high-skilled, with each of them supplying one unit of labor to the market and with labor markets of the two skill levels being segmented. This assumption about skill groups allows us to particularly focus on the low(er) skilled and how their job satisfaction is affected when production moves offshore as supposedly it is them, who bear the brunt of this process. Job satisfaction of all individuals, however, is assumed to depend on consumption possibilities of goods (and thus on income) and on the work itself with the amalgamation subjective and only accessible via introspection. Any explanation of job satisfaction thus has to cater to the economic as well as to the socio-psychological component. Via the budget constraint, both, the semi-pecuniary and the non-pecuniary aspects, are intertwined though as the conjoint effect of preferences and (endogenously determined) incomes shape the extent of offshoring, which in turn is at our focus with respect to job satisfaction.

There is much indication (e.g., Dluhosch and Hens), that both skill levels are affected very differently as to the income effects of offshoring. This is, inter alia, because offshoring, that is, the management of a geographically dispersed production process, requires a larger amount of skilled management services and information and communication technologies. The increase in the demand for skilled management tends to drive up skilled wages, whereas the cost-saving aspect of offshoring proper tends to depress wages of the less skilled. Their wages decline in particular with reference to high-skill-intensive goods as those goods become relatively more expensive.

However important, there is more to offshoring than just income effects which in turn influence consumption possibilities and thus map into (subjective) well-being and job satisfaction in particular. Recall that to capture the subjective part, including non-pecuniary aspects of the workplace as individually perceived, we started out with the notion that job satisfaction is a composite of two things, namely preferences over consumption possibilities (i.e., real income) and socio-psychological aspects related to offshoring. The reason why it is necessary to also track down income effects when exploring job satisfaction properly is that the consumption possibilities provided by the income effects affect output and the extent (and the actual nature) of offshoring, which crucially impacts the workplace.

The composite assumption with respect to job satisfaction is in line with much of the research on workers' perceptions, for instance, in Card et al. (2012) or Lü et al. (2012). However, there, the focus is on inequality. Here, by contrast, we track perceptions as to offshoring as an amalgam of consumption possibilities (stemming from income) and outsourcing-offshoring-related workplace characteristics, both of which are subject to change as competition intensifies by trade being liberalized. The socio-psychological nature is thus accounted for by three components: the individual preference over semi-pecuniary (income) and non-pecuniary (the work itself) elements and how the two components are perceived separately. Notably, we do not exactly specify how those elements are subjectively weighted. In any case, even the non-pecuniary aspects of the work itself are not independent of any income effects related to offshoring. This is because, via affecting demand, they feed back on the extent of offshoring, which in turn affects job satisfaction. Therefore, any socio-economic analysis of offshoring as it relates to job satisfaction must track the income channels carefully.

Income effects indicate that the low(er) skilled most likely suffer from an income-related decline in job satisfaction. Taking into account that socio-psychological aspects of the workplace also change in the course of this process yields an even more diverse picture. For in this case, the low-skilled need not be worse off. More specifically, self-esteem of workers that remain after a wave of offshoring may increase as they feel more important for the company's core business. Offshoring could come along with specific opportunities for career development. The employee might feel valued when belonging to the group of people that faces an increased responsibility after restructuring the company's workflows. Remaining employees may also be more satisfied with their work if tasks are substituted by imported intermediates that are particularly unpleasant. Those additional aspects, next to the pure economic effects on wages, income, and, consequently, consumption, play an important role in individual job satisfaction. They involve a socio-psychological dimension that is important for subjective well-being. If, for example, uninteresting, monotonous, repetitive, or administrative tasks are offshored abroad (or even wiped out completely by the whole offshoring-related restructuring in production), the remaining tasks may be felt as more inspiring and attractive. Consequently, the "new" job might well be considered more interesting. It is also likely that tasks are offshored and in this course rationalized that previously were particularly error-prone, to the effect that jobs without

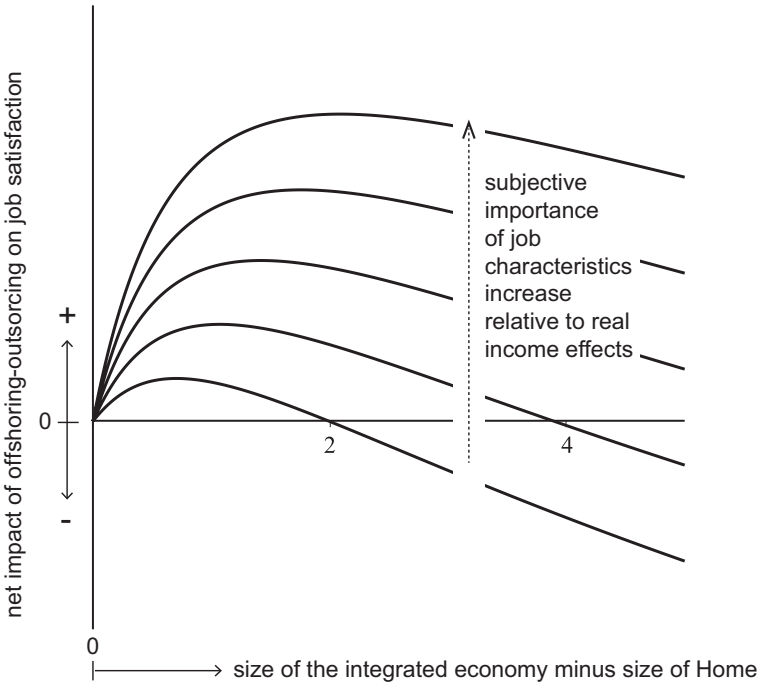

Fig. 1 Net effect of offshoring-outsourcing on individual job satisfaction

those tasks are then be perceived as less stressful. There is a variety of possibilities how the change in job characteristics that comes along with offshoring activities affects the psychological sentiment of those individuals that are still employed and on how satisfied they are with their job.

Figure 1 displays net effects of the composite on job satisfaction for a number of parameter values, which are supposed to capture the subjective character of the issue, including the relative importance of job characteristics as compared to income. Each curve in Fig. 1 shows how both effects of offshoring-outsourcing are amalgamated into overall job satisfaction, with the (relative) weight of workplace characteristics increasing from the bottommost to the topmost curve and with income effects in any case negative.

Net effects on job satisfaction are not positive throughout, but for a range of parameter values. The range of values for which net effects are positive intrinsically depends on the size of the integrated economy. As the integrated economy becomes larger due to advances in information and

communication technologies and the decline in transportation costs, effects of any non-pecuniary benefits tend to become smaller while effects of income at the margin become larger, so that, eventually, negative effects dominate.

In Fig. 1 the tipping point depends on the relative importance of workplace characteristics. The tipping point moves further to the right the more important the workplace characteristics for any given income effect. However, this does not necessarily imply that job satisfaction is certainly to decline at some point in time (provided trade integration continues and deepens). This is because curves may well shift, for instance, with productivity gains related to advances in information and communication technologies. Other influences are preferences over the goods and services and their skill intensity, and the world's composition of the labor force with reference to skills.

Results of our theoretical framework thus have shown that offshoring-related job satisfaction is not only a matter of real income effects, but also a matter of how the work itself is rated, and how both aspects are intertwined. The offshoring-associated change in tasks opens up further socio-psychological dimensions that are important in how individuals perceive their job. The previous dissection of the various influences also provides some indication that this latter effect may (partially) compensate for negative income effects. In the next section, we will approach the issue of job satisfaction from the empirical side in order to get a better grasp on the actual net effects, thus complementing the theoretical explanation above. To serve as a guideline for the empirical section, we condense the key findings of the theoretical model in two empirically testable hypotheses.

Hypothesis 1 Offshoring activities may improve job satisfaction of employees, even when accounting for (possibly negative) income effects.

Hypothesis 2 Offshoring induces a change in tasks. Employees are assumed to be more satisfied with their work if imported intermediates substitute tasks that are particularly unpleasant.

3 Approaching the Effects of Offshoring on Job Satisfaction Empirically

To empirically investigate offshoring effects on individual job satisfaction, it is necessary to combine individual information (e.g., perceived job satisfaction) with more aggregated country-level data (e.g., offshoring activities). In addition to offshoring effects, the empirical section will consider

also other aspects of globalization, namely, aggregated trade flows, trade policy, and outward FDI and thus trace how they feed into individual perceptions about job satisfaction.

Data

Of utmost importance for this contribution are, as outlined in the hypotheses, the effects of offshoring on job satisfaction. Job satisfaction and also a variety of exogenous variables enter the analysis as subjective and not as objective measurable information. Information for the endogenous variable job satisfaction is taken from the International Social Survey Program (ISSP), which, in its 2005 version (Work Orientation III), especially surveys individual perceptions at work. It provides ordinal information on self-reported job satisfaction. Individuals can respond to the question "How satisfied are you in your (main) job?" with "completely satisfied", "very satisfied", "fairly satisfied", "neither satisfied nor dissatisfied", "fairly dissatisfied", "very dissatisfied", or "completely dissatisfied". Thus, the endogenous variable job satisfaction has an ordinal scale with values ranging from 1 ("completely dissatisfied") to 7 ("completely satisfied"). The ISSP provides information on 24,619 individuals in 31 countries.[7]

The exogenous variables of main interest capture different dimensions of globalization at the country level:

1. Trade Policy: we include the trade freedom index by the Heritage Foundation as a proxy for how restrictive trade policies are. Measuring trade policy empirically is not an easy task. Using only particular trade policy measures, as, for example, the weighted average tariff rate, does not account for the heterogeneity in trade policy. It neglects, for example, a wide range of non-tariff barriers. Two frequently cited indices (provided by the Heritage Foundation and the Fraser Institute respectively) are interesting alternatives. Both indices are calculated annually with documented calculation schemes. While the index of the Fraser Institute also captures black market issues and capital flow restrictions, the index of the Heritage Foundation is closer to pure

[7] Table 7 in the Appendix lists all countries covered by the ISSP information (Column 1). Table 8 provides descriptive statistics of all variables used. The distribution of the variable job satisfaction is presented in Table 9. The data can be downloaded at http://www.issp.org. Accessed March 30, 2017.

trade restrictions: it measures the absence of tariff and non-tariff barriers and varies between 0 and 100 on a percentage scale. Tariffs are aggregated using the trade-weighted average tariff rate. A penalty scheme considers additional non-tariff barriers. To test the robustness of our results, we additionally apply the KOF globalization index.[8]

2. Trade Flows: in addition to trade policy, we include information on trade flows. The aforementioned trade policy indices capture more of the option value of international trade, that is, the possibilities opened up by trade liberalization, from which actual trade flows may differ substantially. To take these differences into account, we include actual trade flows as an aggregate measure of openness (that is conventionally the ratio of imports plus exports to GDP at the country level) from the Penn World Tables (PWT).[9]
3. Offshoring and FDI: besides trade policy and trade flows, we include net FDI outflows (over GDP) and offshoring activities as more specific facets of globalization. Information on FDI flows is obtained from the World Development Indicators (WDI). Offshoring is calculated using Eurostat input-output tables.[10] Applying the calculation schemes presented in Horgos (2009) and Castellani et al. (2013), we consider a measure of import penetration as aggregated offshoring index (ratio of imported to domestically produced inputs), material offshoring (ratio of imported to domestically produced material inputs), and service offshoring (ratio of imported to domestically produced service inputs). As has, inter alia, been shown in these recent contributions, import penetration is a good proxy for offshoring since it directly measures the economic activity that is at the very heart of the offshoring phenomenon, namely, the use of imported inputs instead of domestically sourced intermediates. In line with the majority of recent contributions to

[8] For more detailed information on how the index of the Heritage Foundation is calculated or to download the index, see http://www.heritage.org. Accessed March 30, 2017. To download the index of the Fraser Institute, see http://www.freetheworld.com. Accessed March 30, 2017. For further information about the KOF globalization index, see http://globalization.kof.ethz.ch. Accessed March 30, 2017.

[9] The PWT data set is available at http://www.rug.nl/ggdc/productivity/pwt/. Accessed March 30, 2017. This contribution uses version 7.0 available since June 2011.

[10] The WDI are available at http://data.worldbank.org/data-catalog/world-development-indicators. Accessed March 30, 2017. We use the WDI and GDP data from April 19, 2012. The input-output tables from Eurostat are available at http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables. Accessed March 30, 2017.

international economics, the offshoring index as it is employed here does not distinguish between the ownership of foreign productions: it contains imports of inputs at arms' length as well as imports from foreign subsidiaries. As will be seen shortly, interesting results emerge inter alia with respect to the interplay between FDI outflows and offshoring. When additionally controlling for offshoring activities that capture international trade flows related to the relocation of production, pure investment decisions are left for the FDI variable. Both of the components affect individual job satisfaction quite differently.

When examining the implications of the different dimensions of globalization on individual job satisfaction, it is necessary to include individual as well as country-level control variables.[11] As illuminated by theory, income effects are an important determinant to assess job satisfaction. For the empirical strategy it is thus necessary to adequately control for individual income. We therefore added individual income as ordered information to the set of control variables.[12] In addition, we consider job security (as ordered variable with five categories ranging from "strongly disagree" (namely that the job is considered secure) to "strongly agree"), gender, age, age-square, marital status (married, widowed, divorced, single, and "married, but living separately" as reference category), and years of schooling. For reasons of consistency, all individual control variables are taken from the ISSP 2005.

In contrast to many contributions to the well-being literature that draw on more aggregated measures like happiness or life satisfaction, this contribution shifts the focus on job satisfaction in particular. Therefore, variables that capture different job characteristics are of additional interest. As to these job characteristics, the ISSP provides information on whether employees rate their job as interesting, independent, exhausting, hard work, and/or stressful. Question 10 of the ISSP asks employees if they (1) strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, or (5) strongly disagree with the statements that their (main) job is interesting or that they can work independently. Question 11 asks how often

[11] As to the control variables which are to be used for the regressions in this contribution, we refer to Scheve and Slaughter (2001, 2004), Hanson et al. (2007), Bjørnskov et al. (2008), Skalli et al. (2008), Geishecker (2008, 2012), Geishecker et al. (2012).

[12] Because we obtain our individual variables from ISSP survey data, we have ordinal information on income. Individuals are asked if they rate their income as high. Interviewees have five options, ranging from "strongly disagree" to "strongly agree".

employees come home from work exhaustive, do have to do hard physical work, or do find their work stressful. The interviewed persons can also choose from five different answers: (1) always, (2) often, (3) sometimes, (4) hardly ever, or (5) never. We use this information to define ordinally scaled job-characteristic variables with five categories. We will deepen the discussion on the different job characteristics in the following section.

In order to account for macroeconomic income effects on individual job satisfaction, we include the log of GDP per capita (in international dollar, PPP converted at 2005 constant prices, obtained from PWT). As additional macroeconomic control variables we consider the log of population (also obtained from the PWT). Overall, the data consists of cross-section information on 24,619 individuals from 31 countries in the year 2005.

Estimation and Results

To examine how offshoring affects individual job satisfaction, we apply a cross-section analysis for 2005. Because the endogenous information (individual job satisfaction) has an ordinal scale whereas the main regressors (trade policy, trade flows, and the more specific dimensions of trade flows, namely, FDI and offshoring) are metric variables with information in percent, we employ an ordered logistic regression model.[13] In estimating the model, we face an additional challenge though in that the endogenous variable is at the level of the individual whereas the exogenous variables of main interest are economy-wide aggregates. One common solution to this problem is to control for enough individual and country-level variance and to cluster the errors at the country level (see Chamberlain, 1980; Alesina et al., 2004; or Ferrer-i-Carbonell & Frijters, 2004). We thus derive the ordered logistic regression model from the latent variable estimation

$$JS_{ij}^{*} = \beta_1 TP_j + \beta_2 TF_j + \beta_3 FDI_j + \beta_4 OFF_j^z + \beta_5 IC_{ij} + \beta_6 CC_j + \in_{ij}$$

[13] The additional control variables are differently scaled. We combine dummy variables, metric variables, and variables with ordinal scales. From a purely statistical angle, all exogenous variables in ordered logit models are assumed to have a metric scale. Estimation results with ordered variables on the right-hand side are, however, very similar concerning size and significance of the estimated coefficients. Nevertheless, we will come back to this issue when testing the robustness of the empirical model below.

with vector JS_{ij}^{*} containing latent job satisfaction of individual i living in country j. The seven ordered categories of job satisfaction observed are defined as $JS_{ij}=1$ if $JS_{ij}^{*} \leq a_1$, $JS_{ij}=2$ if $a_1 < JS_{ij}^{*} \leq a_2$, ..., $JS_{ij}=6$ if $a_5 < JS_{ij}^{*} \leq a_6$, $JS_{ij}=7$ if $a_6 < JS_{ij}^{*}$. The four exogenous variables of main interest are trade policy, TP_j (proxied by the trade-freedom index from the Heritage Foundation), trade flows, TF_j (that is the openness measure), FDI net outflows, FDI_j, and offshoring activities, OFF_j^z, all measured at the level of country j. We consider three different variants z of offshoring: aggregated ($z=a$), material ($z=m$), and service offshoring ($z=s$). Individual control variables are combined in matrix IC_{ij}; matrix CC_j captures the macroeconomic information. Because we are interested in the implications of aggregate variables in microeconometric regressions, we cluster the errors (ϵ_{ij}) at the country level. To achieve robust standard errors, the Huber/White sandwich estimator is applied in all of the regressions.

One serious problem in these kinds of regressions is multicollinearity. If multicollinearity characterizes the data, standard errors for the coefficients tend to be inflated. To test for multicollinearity, we calculate the variance inflation factor (VIF). The VIF measures the strength of the interrelation of the exogenous variables, indicating how much of the inflation of the standard error is driven by collinearity. A common benchmark to identify multicollinearity is a VIF measure above 10. The VIF measures of the variables in our estimation are, however, far below this benchmark.[14]

Table 1 presents first estimation results, lending support to Hypothesis 1.[15] Column (1) adds our four main exogenous variables to the set of individual and country-level control variables. The sample in Table 1 refers to European economies because calculations of offshoring indices are based on Eurostat input-output tables. We thus end up with 8863 individuals in

[14] The only exceptions are age and age-square. A strong correlation between those variables, however, is expected, and, as they are included for control reasons only, not biasing the estimates of the coefficients of main interest. Regarding Column 1 of Table 1, the calculated VIF measures for the macroeconomic control variables are 2.90 (Trade Freedom), 5.09 (Openness), 4.44 (Outward FDI), and 5.70 (Offshoring).

[15] All tables in this section present results for the exogenous variables of main interest. Notably, size and significance of the control variables are similar in all of the estimated models. Qua example, Table 10 in the Appendix presents estimation results of the full set of variables, including those for control reasons and corresponding to the results presented in Table 1, Column (1). Additional tables presenting the full set of regression results of specific models can be obtained from the authors upon request.

Table 1 Effects of international trade on individual job satisfaction

	(1)	*(2)*	*(3, odds ratios)*	*(4, odds ratios)*
Endogenous variable: individual job satisfaction				
Trade freedom	-0.0437*	-0.0697***	0.9572*	0.9327***
	(0.0265)	(0.0175)	(0.0253)	(0.0163)
Openness	-0.0080***	-0.0033	0.9920***	0.9967
	(0.0027)	(0.0030)	(0.0027)	(0.0030)
Outward FDI	-0.0342	-0.0470**	0.9664	0.9541**
	(0.0286)	(0.0232)	(0.0277)	(0.0221)
Offshoring	1.3731***	–	3.9476***	–
	(0.3830)		(1.5118)	
Material offshoring	–	-0.2725	–	0.7615
		(0.3176)		(0.2419)
Service offshoring	–	1.0385***	–	2.8248***
		(0.2205)		(0.6229)
Ind. controls	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes
Obs.	8863	8863	8863	8863
Clusters	12	12	12	12
Pseudo R2	0.1359	0.1368	0.1359	0.1368
Wald Chi2	890***	1761***	890***	1761***

Robust standard errors in parentheses; *, **, *** Statistical significance at the level of 10%, 5%, 1%

12 European countries.[16] In these European economies, a more liberal trade policy (trade freedom) is significantly correlated with lower levels of individual job satisfaction. Job satisfaction is also lower in countries that have a larger openness index, with the nexus statistically significant at the level of 1%. However, the more intermediates are sourced abroad (which implies a high offshoring index), the higher is the job satisfaction as stated by the individuals of that country. The effect is statistically significant at the level of 1%. The pseudo R-square of 0.1359 and the significant Wald test resulting in a chi-squared value of 890 indicate that the overall model is fitted well.

Figure 2 displays (average) marginal effects of an increase in the total import penetration (i.e., the offshoring index) at different values of the import penetration ratio with all other covariates as observed. The import penetration ranging from 20 to 100% covers approximately the whole

[16] Table 7 in the Appendix provides a synopsis as to the countries covered by the data set and how much information is lost when controlling for the different variables.

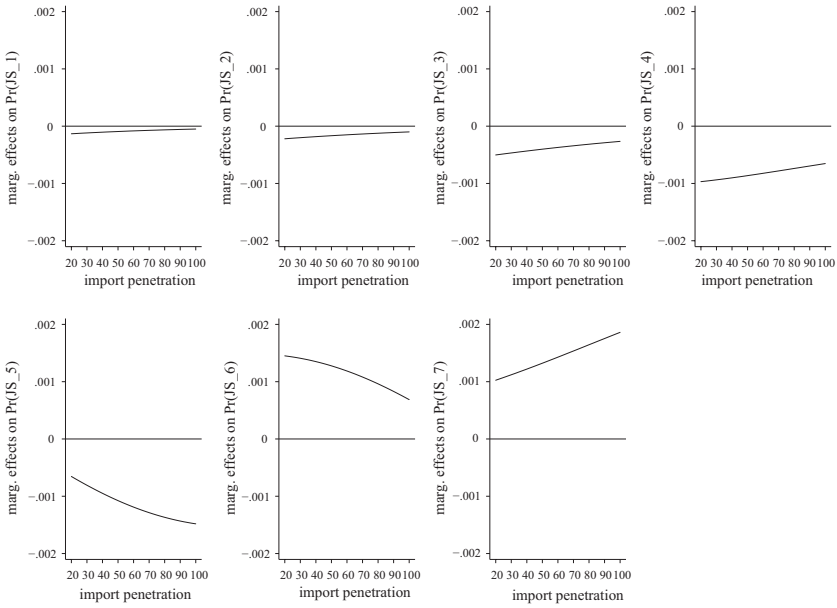

Fig. 2 Marginal effects of import penetration for different job satisfaction scores

domain of the variable in the data set. The first row in Fig. 2 shows scores 1–4 in self-reported job satisfaction, starting with the bottommost score (*JS*_1); the second row shows scores 5–7, thus ending with the topmost score (*JS*_7). The disaggregated perspective according to the level of job satisfaction substantiates results of Table 1 further: most notably, marginal effects are negative at the lower end of the satisfaction scale and positive at the upper end. This means that predicted probabilities of low job satisfaction decrease as the import penetration increases while those of high(er) job satisfaction increase. The pattern applies to the whole domain of import penetration ratios covered in the data: individuals tend to report higher job satisfaction at higher levels of offshoring.

Switching perspective to job satisfaction and the quality of work life thus delivers interesting insights beyond the purely pecuniary consequences of globalization in terms of incomes and their distribution. Obviously, and contrary to conventional beliefs, offshoring has a positive effect on job satisfaction. It is important though to keep in mind that our analysis focuses on individuals that are still employed after offshoring has

taken place. This enables us to examine how working conditions differ for various levels of integration.

In the regression of Column (2) in Table 1, we distinguish in addition between material and service offshoring. Results show that import-penetration in services has positive explanatory power. The estimated coefficient is positive and significant at the level of 1%. In countries where relative more service inputs are purchased from abroad, employees report higher job satisfaction. Outward FDI shows a negative correlation with individual job satisfaction that also shows up statistically significant at the level of 5% when estimating the model presented in Column (2). As indicated by the pseudo R-square and the Wald test, the goodness of fit increases when switching to this specification. Calculating the odds ratios of the logit model reveals that offshoring (Column 3), as well as service offshoring (Column 4) in particular, is substantially correlated with job satisfaction.

Overall, results support Hypothesis 1: individual job satisfaction is significantly higher in countries with extensive offshoring. And, notably, this result holds after controlling inter alia for individual income. By contrast, in countries with relative high net outflows of FDI, individual job satisfaction is significantly lower. The reason for the FDI effect seems straightforward: capital invested abroad means that there is, ceteris paribus, less capital at home which might have otherwise, for example, improved domestic job or working conditions. The offshoring effect, however, requires additional explanation.

Estimating the effect of offshoring on job satisfaction is not straightforward, especially not when also trying to account for the role of income and socio-psychological determinants. As alluded to in the theoretical section, offshoring has inter alia an effect on job satisfaction because it induces also non-pecuniary effects beyond the changes in income. Accounting for this in an empirical analysis is not straightforward since we are not able to empirically observe all of the manifold socio-psychological aspects. However, as pointed out while introducing the data, we can get a little bit closer to the heart of the matter as the ISSP data provides information on different job characteristics.

Table 2 compares various estimation results regarding the nexus between offshoring and job satisfaction. The estimations vary as we include or drop the control variables for income and/or different job characteristics. We are thus able to disentangle various channels. When controlling for income and job characteristics, the offshoring variable still captures

Table 2 Effects of international trade on individual job satisfaction (with and without controls for income and job characteristics)

	(1)	*(2)*	*(3)*	*(4)*
Endogenous variable: individual job satisfaction				
Contr. income	Yes	No	Yes	No
Contr. job charact.	Yes	No	No	Yes
Trade freedom	-0.0437*	0.0063	-0.0134	-0.0300
	(0.0265)	(0.0384)	(0.0406)	(0.0238)
Openness	-0.0080***	-0.0047	-0.0077**	-0.0060**
	(0.0027)	(0.0038)	(0.0035)	(0.0026)
Outward FDI	-0.0342	-0.0790*	-0.0581	-0.0455
	(0.0286)	(0.0438)	(0.0400)	(0.0292)
Offshoring	1.3731***	1.4736***	1.6374***	1.2507***
	(0.3830)	(0.3285)	(0.3037)	(0.3483)
Ind. controls	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes
Obs.	8863	9028	8976	8901
Clusters	12	12	12	12
Pseudo R2	0.1359	0.0255	0.0454	0.1263
Wald Chi2	890***	2643***	2022***	2077.82***

Robust standard errors in parentheses; *, **, *** Statistical significance at the level of 10%, 5%, 1%

socio-psychological aspects that we are not able to directly observe and thus control for. Accordingly, the correlation between offshoring and job satisfaction is positive and statistically significant at the level of 1% in all four model variants (Columns 1–4). Thus, non-pecuniary, socio-psychological aspects that are carried by the offshoring variable turn out to be important determinants of job satisfaction. To name just a few of these aspects, the trust of being important for the company might increase for those individuals that are still employed after offshoring activities. The remaining employees might also have new career-development opportunities after production parts have moved offshore. This may be the case if responsibilities increase in tandem with the change in tasks for the remaining employees. The socio-psychological aspects that are carried by the offshoring variable after controlling for individual income and job characteristics account for the statistically significant positive correlation between offshoring and job satisfaction (see Column 1). Due to the ambiguous interaction between offshoring and individual income, the positive correlation between the offshoring variable and job satisfaction is smaller in size and statistical significance when not controlling for income

specifically (Column 4). If we do not control for different job characteristics (Columns 2 and 3), the size of the positive correlation between offshoring and job satisfaction is increasing, whereas the explaining power of the overall model (pseudo R2) decreases dramatically. This points toward an important role and additional positive correlation between offshoring-relevant workplace characteristics and job satisfaction.

Results thus far immediately raise the question as to which workplace characteristics are precisely affected by offshoring. A closer look at this question helps to identify the exact channels via which offshoring influences perceptions of the work itself, the workplace, and thus job satisfaction. ISSP data allows to examine the question more closely. It provides information as to whether individuals perceive their job as interesting, independent, exhausting, hard work, or stressful. In the next set of regressions, we will shift the focus of the analysis to a more disaggregated level and examine these diverse facets of job satisfaction in greater detail.

Before estimating the effects of offshoring on these (subjective) workplace characteristics, Table 3 displays the correlation matrix of the different job characteristics. The pairwise correlations in Table 3 show that job satisfaction is positively correlated with the job being perceived as interesting and/or independent. By contrast, a job which is perceived as exhausting, hard work, or stressful is negatively correlated with individual job satisfaction. In the next set of regressions we use the separate job characteristics as different endogenous variables instead of the aggregate

Table 3 Correlation matrix of job characteristics

	Job satisfaction	*Interesting*	*Independent*
Job satisfaction	1		
Interesting	0.4795*	1	
Independent	0.2684*	0.4014*	1
Exhausting	-0.1662*	-0.0946*	-0.0729*
Hard work	-0.1210*	-0.1599*	-0.0761*
Stressful	-0.1873*	-0.0390*	-0.0525*
	Exhausting	*Hard work*	*Stressful*
Job satisfaction			
Interesting			
Independent			
Exhausting	1		
Hard work	0.3758*	1	
Stressful	0.4453*	0.1821*	1

*Significant pairwise correlation coefficient at the level of 1%

measure of individual job satisfaction. This procedure enables us to find out which characteristic of the job is affected the most by offshoring and thus lends additional support to Hypothesis 2. Table 4 presents the results.

Results show that the international relocation of parts of production (that is high values of the offshoring index) correlates positively with the jobs being considered more interesting (Column 1). The effect is statistically significant at the level of 5%. Thus, in economies with high levels of offshoring, employees perceive their job as relatively more interesting. In addition, offshoring shows a statistically significant negative impact on the stressfulness of tasks (Column 5). However, all that glitters is not gold: even if jobs are reported to be less stressful, individuals evidently perceive their tasks to be more exhausting (Column 3).

Table 4 Effects of international trade on different job characteristics

Endogenous variable	*Interesting*	*Independent*	*Exhausting*
	(1)	*(2)*	*(3)*
Trade freedom	0.1318*** (0.0431)	-0.1004** (0.0419)	0.0573 (0.0418)
Openness	-0.0096*** (0.0038)	0.0031 (0.0066)	-0.0101*** (0.0027)
Outward FDI	0.0302 (0.0433)	-0.0916 (0.0784)	-0.0253 (0.0579)
Offshoring	2.0456*** (0.3202)	-0.3640 (1.1870)	2.6367*** (0.6693)
Ind. controls	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes
Obs.	8927	8927	8972
Clusters	12	12	12
Pseudo R2	0.1271	0.1138	0.1492
Wald Chi2	2215***	619***	668***
Endogenous variable	*Hard work*	*Stressful*	
	(4)	*(5)*	
Trade freedom	-0.1093*** (0.0245)	-0.0275 (0.0203)	
Openness	-0.0025 (0.0023)	0.0052*** (0.0017)	
Outward FDI	0.0719* (0.0387)	0.0288 (0.0313)	
Offshoring	-0.4468 (0.3434)	-2.0509*** (0.2570)	
Ind. controls	Yes	Yes	
Macro controls	Yes	Yes	
Obs.	8927	8927	
Clusters	12	12	
Pseudo R2	0.0897	0.1005	
Wald Chi2	1435***	1226***	

Robust standard errors in parentheses

*, **, *** Statistical significance at the level of 10%, 5%, 1%

4 Picking Out the Cherries or (Simply) Statistical Cherry-Picking?

Job satisfaction data thus yield novel insights into the implications of offshoring, reaching substantially beyond the ones on income and distributional consequences. To check how robust these effects are, we apply a variety of robustness tests on the empirical setup.

Although the theoretical framework has shown that income effects need not in any case be negative, it also showed that low- and high-skilled individuals might well be affected differently. Therefore, we address issues of skill also empirically by differentiating between employees with low and high education and, additionally, between employees with low and high income. Table 5 has the results.

A comparison of the results presented in Columns (1) and (2) (i.e., distinguishing by years of education) reveals that the positive correlation of offshoring with individual job satisfaction holds for both skill groups. However, estimated odds ratios are higher for low-educated individuals.

The same result shows up when using income as distinguishing variable (Columns 3 and 4): in both regressions, offshoring turns out to have a significant positive effect on job satisfaction. Estimated odds ratios are, however, higher for individuals that state that their income is relative low. In light of our findings that jobs are perceived as being more interesting, more exhausting, but less stressful in countries with high offshoring indices, the group of the low skilled and those with low income can be assumed to have a stronger leverage. Notably, this empirical result is in line with our theoretical explanation on how offshoring affects perceptions of the lower skilled in particular. Moreover, this result is of special interest in light of findings illuminating negative employment and income effects that offshoring might impose on low-skilled labor. The lower skilled seem to be partly compensated as those still employed report relative higher job satisfaction when offshoring takes place.

In checking further the robustness of the estimations, we replace the trade freedom index provided by the Heritage Foundation with two alternative measures. Because the empirical measurement of trade policy is intensively discussed in the economic literature (see, e.g., Anderson and Neary 1994; Rose 2004; Dluhosch and Horgos 2013), it is necessary to test whether our results also hold when including alternative proxies for trade policy. Notably, our results are invariant to the particular trade freedom index that is being used: Table 6 displays results when using the trade

Table 5 Effects on job satisfaction, distinguishing between different groups of labor

Endogenous variable: individual job satisfaction; estimated odds ratios				
	Education (years)		*Income*	
	Low	*High*	*Low*	*High*
	(1)	*(2)*	*(3)*	*(4)*
Trade freedom	0.9538* (0.0247)	0.9673 (0.0294)	1.0297 (0.0196)	0.9177** (0.0351)
Openness	0.9889*** (0.0036)	0.9928*** (0.0024)	0.9909*** (0.0024)	0.9924** (0.0039)
Outward FDI	0.9849 (0.0300)	0.9647 (0.0266)	0.9631 (0.0306)	0.9651 (0.0323)
Offshoring	5.1641*** (3.0362)	3.9430*** (1.3902)	5.8383*** (2.4297)	2.6445** (1.2373)
Ind. controls	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes
Obs.	3190	5673	2179	4061
Clusters	12	12	12	12
Pseudo R2	0.1143	0.1548	0.1083	0.1257
Wald Chi2	306***	891***	1311***	395***

Robust standard errors in parentheses; *, **, *** Significant at 10%, 5%, 1%

Education distinguishes by years of education: (Column 1) considers individuals with 0–11 years, (2) individuals with more than 11 years

Income distinguishes according to the following classification

(3) refers to individuals that disagree or strongly disagree that their income is high

(4) refers to individuals that agree or strongly agree that their income is high

freedom index provided by the Fraser Institute (Columns 1 and 2) or the KOF globalization index (Columns 3 and 4) as alternative proxies for liberal trade policy.

Offshoring is in any case positively correlated with individual job satisfaction and at high levels of statistical significance. However, when using the two alternative measures, results of the outward FDI variable are no longer statistically significant. The reason for this finding is the structure of the two alternative measures. Both of them include restrictions on capital flows, which grab variance from the outward FDI variable.

To prove that results do not primarily depend on the specific logit estimator that is used for the regressions presented above, we additionally

Table 6 Effects of international trade on individual job satisfaction using alternative trade-policy indices

Endogenous variable: individual job satisfaction				
	(1)	*(2)*	*(3)*	*(4)*
Index Fraser Institute	-0.0490	-0.2414	–	–
	(0.1660)	(0.2159)		
KOF globalization index	–	–	-0.0055	-0.0246**
			(0.0045)	(0.0101)
Openness	-0.0080**	0.0007	-0.0082***	0.0022
	(0.0034)	(0.0055)	(0.0029)	(0.0045)
Outward FDI	-0.0201	-0.0273	-0.0181	-0.0154
	(0.0378)	(0.0289)	(0.0399)	(0.0292)
Offshoring	1.7225***	–	1.7244***	–
	(0.3202)		(0.2997)	
Material offshoring	–	-0.2816	–	-0.6597
		(0.5041)		(0.4510)
Service offshoring	–	1.3574***	–	1.607***
		(0.4208)		(0.3405)
Ind. controls	Yes	Yes	Yes	Yes
Macro controls	Yes	Yes	Yes	Yes
Obs.	8863	8863	8863	8863
Clusters	12	12	12	12
Pseudo R2	0.1355	0.1359	0.1355	0.1367
Wald Chi2	1105***	5871***	1103***	5978***

Robust standard errors in parentheses; *, **, *** Statistical significance at the level of 10%, 5%, 1%

estimate the whole set of regressions using probit estimations. Here, as well, results lend support to our hypotheses and the initial findings presented above. High offshoring activities of a country correlate positively with job satisfaction as perceived by employees. Probit results also show that the main effect is rooted in service offshoring.

Finally, several of the individual control variables are measured at an ordinal scale. Although this is not significantly biasing results, prima facie, this does not square well with the fact that the exogenous variables in logit and probit models are assumed to be metric. Therefore, a final set of robustness checks replaces the ordered exogenous variables with their (metric) dummy-variable counterparts. Results are robust.[17]

[17] Tables for the additional robustness checks (as to the probit estimates and metric counterparts) can be obtained from the authors upon request.

5 Conclusions

Previous research on the globalization of production and labor almost exclusively focused on wages and distributional consequences for skill levels, occupations, sectors, industries, or even firms. Surprisingly less information is on the table as to the psychological consequences and various indicators of subjective well-being. This deficiency is all the more striking since psychological and sociological research has shown that subjective well-being need not move in tandem with incomes. The question thus arises as to how job satisfaction is affected, in particular with respect to those whose tasks are more easily susceptible to offshoring and to restructuring and who might be subject to adverse income effects. On the face of it, the answer seems to be straightforward: accordingly, low-skilled labor suffers a decline in job satisfaction in the process of offshoring—not the least because offshoring is thought to have a bias toward low-skilled labor and thought to diminish income perspectives.

However, this notion does not stand the test. We show that the picture is much more diverse with the various dimensions of international competition affecting subjective well-being quite differently. While some indicators of international trade support conventional beliefs, our results show that job satisfaction is significantly higher in countries with relative high offshoring activities. And, we are able to trace these subjectively positive effects of offshoring back to changes in job characteristics in particular. Beyond job characteristics, however, offshoring also comes along with other socio-psychological aspects that we are not able to control for, for instance, trust, self-categorization, knowledge, or the opportunity for better career development. By illuminating non-pecuniary aspects thus far being underrated, our results suggest that future research may also deliver interesting results by further elaborating on the possible effects of different non-pecuniary aspects of offshoring.

In effect, our approach is novel in two respects. First, we present a framework that is capable of capturing and thus explaining job satisfaction in conjunction with the income and distributional effects of offshoring and the globalization of production. These theoretical considerations indeed suggest that even those who remain employed but suffer from intensifying competition from abroad and from more tasks being offshored may actually be more satisfied with their work. The effect is driven by a socio-psychological dimension that shows up when examining determinants of job satisfaction. With offshoring activities, tasks change. Employees are more satisfied with their work if unpleasant tasks are

relocated abroad. Those remaining employed thus score higher on the job satisfaction scale. Second, we take our theoretical results to the data. As to our main variables, job satisfaction and offshoring activities, we use information from the ISSP and the European IO Database. And, we tap the PWT and the World Bank's WDI, in an attempt to control for variables other than offshoring that might influence job satisfaction. Running a cross-section ordered logistic regression model that combines this information lends support to our theoretical results. Accordingly, job satisfaction is on average rated higher in countries with comparatively high offshoring activities, inter alia because of significant changes in job characteristics. This result proves to be robust in several variants of the empirical setup and the data. Hence, accounting for subjective well-being yields new insights and opens up substantially different perspectives on offshoring.

The approach lends itself to a number of extensions. One question is certainly whether there is a notable difference between offshoring in general and trade within GVCs in particular. Their competitive effects may differ, and so may their impact on job satisfaction. Established links, which are already embedded into existing production networks, such as in GVCs, may more easily accommodate shocks and shifts in the geographical distribution of production which in turn may affect job satisfaction. And, other than offshoring per se, GVCs may strengthen complementarities as opposed to substitution effects. Another promising route for further research is the multifaceted nature of job satisfaction and the seemingly different situation as to the various components displayed, for instance, in (East) Asian and Western data and whether those differences find any correspondence in trade data. Trying to identify geographical, cultural, organizational, or institutional differences may deliver additional information on important intervening variables or even policies in the nexus between international competition and job satisfaction. The ISSP data meanwhile provides four surveys focusing on work orientation (1989, 1997, 2005, and 2015). Interesting results may emerge when exploiting a possible panel-data dimension and investigating cross-country variations. This may also establish new links and novel insights to development economics and the respective literature.

Acknowledgments This paper very much benefited from discussions in a number of seminars and conferences and from the comments and suggestions of anonymous referees of this journal for which we are very grateful.

Appendix

Table 7 List of countries covered by the data set

Country	*ISSP*	*Macro controls*	*Trade policy and openness*	*FDI*	*Offshoring*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Australia	✓	✓	✓	✓	.
Belgium/Flanders	✓	✓	✓	✓	✓
Bulgaria	✓	✓	✓	✓	.
Canada	✓	✓	✓	✓	.
Cyprus	✓	✓	✓	✓	.
Czech Republic	✓	✓	✓	✓	.
Denmark	✓	✓	✓	✓	✓
Dominican Republic	✓	✓	✓	.	.
Finland	✓	✓	✓	✓	✓
France	✓	✓	✓	✓	✓
Germany	✓	✓	✓	✓	✓
Great Britain	✓	✓	✓	✓	.
Hungary	✓	✓	✓	✓	✓
Ireland	✓	✓	✓	✓	✓
Israel	✓	✓	✓	✓	.
Japan	✓	✓	✓	✓	.
Latvia	✓	✓	✓	✓	.
Mexico	✓	✓	✓	✓	.
New Zealand	✓	✓	✓	✓	.
Norway	✓	✓	✓	✓	✓
Philippines	✓	✓	✓	✓	.
Portugal	✓	✓	✓	✓	✓
Russia	✓	✓	✓	✓	.
Slovenia	✓	✓	✓	✓	✓
South Africa	✓	✓	✓	✓	.
South Korea	✓	✓	✓	✓	.
Spain	✓	✓	✓	✓	✓
Sweden	✓	✓	✓	✓	✓
Switzerland	✓	✓	✓	✓	.
Taiwan	✓	.	✓	.	.
United States	✓	✓	✓	✓	.
Sum	31	30	31	29	12

Data: The data consists of the ISSP 2005: Work Orientation III (all individual variables), the PWT and WDI (macroeconomic control variables), trade policy information from the Heritage Foundation, the Fraser Institute and the KOF Globalization index, Openness (taken from the PWT), and Offshoring information for European countries (calculated from input-output data provided by Eurostat)

Table 8 Descriptive statistics

	Job satisfaction	*Job security*	*High income*	*Interesting job*	*Independent job*
N	24,619	24,439	24,548	24,595	24,565
Mean	5.251026	3.605753	2.732076	3.795528	3.768125
SD	1.204405	1.138506	1.08844	1.002654	1.094412
Variance	1.450592	1.296196	1.184701	1.005315	1.197738
min	1	1	1	1	1
max	7	5	5	5	5
	Exhausting job	*Hard work*	*Stressful job*	*Male*	*Age*
N	24,677	24,647	24,644	43,392	43,224
Mean	2.322608	1.53118	2.133379	0.4564436	45.75669
SD	0.9671773	1.296406	1.06808	0.498105	17.05782
Variance	0.9354319	1.68067	1.140794	0.2481086	290.9693
min	0	0	0	0	15
max	4	4	4	1	98
	Age^2	*Married*	*Widowed*	*Divorced*	*Single*
N	43,224	43,140	43,140	43,140	43,140
Mean	2384.637	0.5753825	0.0785814	0.0665508	0.2585535
SD	1670.906	0.4942905	0.2690874	0.2492453	0.4378448
Variance	2791928	0.2443231	0.072408	0.0621232	0.1917081
min	225	0	0	0	0
max	9604	1	1	1	1
	Education (years)	*GDP p.c.*	*Population*	*Trade freedom*	*Openness*
N	39,639	43,440	43,440	43,440	43,440
Mean	11.71334	9.960031	9.839852	77.77757	80.39412
SD	3.988263	0.6590892	1.347059	4.883895	32.77223
Variance	15.90624	0.4343986	1.814567	23.85244	1074.019
min	0	7.808705	6.918695	63.2	26.07303
max	49	10.79346	12.59728	83.6	156.7588
	Outward FDI	*Offshoring*			
N	39,311	16,350			
Mean	2.535067	0.4207911			
SD	3.543774	0.2085658			
Variance	12.55834	0.0434997			
min	-4.873191	0.2035569			
max	13.65024	1.006923			

Table 9 Distribution of individual job satisfaction

Parameter value	*Freq.*	*Percent*	*Cum.*
Completely dissatisfied	262	1.06	1.06
Very dissatisfied	486	1.97	3.04
Fairly dissatisfied	1285	5.22	8.26
Neither satisfied nor dissatisfied	2677	10.87	19.13
Fairly satisfied	9449	38.38	57.51
Very satisfied	6987	28.38	85.89
Completely satisfied	3473	14.11	100.00

Data: ISSP 2005: Work Orientation III

Table 10 Effects of international trade on individual job satisfaction

Endogenous variable: individual job satisfaction			
full set of variables, including all individual and macroeconomic controls			
	(1)		*(1 cont.)*
Job security	0.1673*** (0.0207)	GDP p.c. (log)	0.5421* (0.3022)
High income	0.3243*** (0.0417)	Population (log)	0.0472 (0.0570)
Interesting job	0.9716*** (0.0755)	Trade freedom	-0.0437* (0.0265)
Independent job	0.1893*** (0.0335)	Openness	-0.0080*** (0.0027)
Exhausting job	-0.2336*** (0.0495)	Outward FDI	-0.0342 (0.0286)
Hard work	0.0430 (0.0295)	Offshoring	1.3731*** (0.3830)
Stressful job	-0.3039*** (0.0397)	Obs	8863
Male	-0.0932* (0.0512)	Clusters	12
Age	-0.0306** (0.0125)	Pseudo R2	0.1359
Age^2	0.0004*** (0.0001)	Wald Chi2	890***
Married	0.1930 (0.1859)		
Widowed	0.3862 (0.3059)		
Divorced	0.0711 (0.2059)		
Single	0.0085 (0.2138)		
Education (years)	-0.0459*** (0.0091)		

Robust standard errors in parentheses; *, **, *** Statistical significance at the level of 10%, 5%, 1%

Reference category of marital status is "married, but living separately"

References

Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and happiness. Are Europeans and Americans Different? *Journal of Public Economics, 88*, 2009–2042.

Amador, J., & Cabral, S., (2016). Global value chains: A survey of drivers and measures. *Journal of Economic Surveys, 30*(2), 278–301.

Anderson, J. E., & Neary, P. J. (1994). Measuring the restrictiveness of trade policy. *World Bank Economic Review, 8*(2), 151–69.

Becker, S. O.,& Muendler, M. (2015). Trade and tasks: An exploration over three decades in Germany. *Economic Policy, 30*(84), 589–641.

Bjørnskov, C., Dreher, A., & Fischer, J. A. V. (2008). Cross-country determinants of life satisfaction: Exploring different determinants across groups in society. *Social Choice and Welfare, 30*(1), 119–73.

Böckerman, P., & Maliranta, M. (2013). Outsourcing, occupational restructuring, and employee well-being: Is there a silver lining? *Industrial Relations, 52*(4), 878–914.

Brandt, L., & Thun, E. (2016). Constructing a ladder for growth: Policy, markets, and industrial upgrading in China. *World Development, 80*(1), 78–95.

Card, D., Mas, A., Moretti, E., & Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review, 102*(6), 2981–3003.

Cazes, S., Hijzen, A., & Saint-Martin, A. (2015). Measuring and assessing job quality: The OECD job quality framework. *OECD Social, Employment and Migration Working Papers*, No. 174, Paris: OECD.

Castellani, D., De Benedictis, L., & Horgos, D. (2013). Can we really trust offshoring indices? *Structural Change and Economic Dynamics, 25*(C), 159–72.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies, 47*(1), 225–38.

Clark, A. E., & Oswald, A. J. (1998). Comparison-concave utility and following behaviour in social and economic settings. *Journal of Public Economics, 70*(1), 133–55.

Clark, A. E., Kristensen, N., & Westergård-Nielsen, N. (2009). Job satisfaction and co-worker wages: Status or signal? *The Economic Journal, 119*(March), 430–47.

Colantone, I., Crinò, R., & Ogliari, L. (2015). The hidden cost of globalization: Import competition and mental distress. *CESifo Working Paper*, No. 5586, Munich: CESifo.

Datta, D. K., Guthrie, J. P., Basuil, D., & Pandey, A. (2010). Causes and effects of employee downsizing: A review and synthesis. *Journal of Management, 36*(1), 281–348.

Diener, E., Tay, L., & Oishi, S. (2013). Rising income and the subjective well-being of nations. *Journal of Personality and Social Psychology, 104*(2), 267–76.

Dluhosch, B., & Hens, T. (2016). A rigorous approach to business services offshoring and North-North trade. *Applied Economics, 48*(15), 1390–1401.

Dluhosch, B., & Horgos, D. (2013). Trading up the happiness ladder? *Social Indicators Research, 113*(3), 973–90.

Easterlin, R. A. (1974). Does economic growth improve the human lot? In P. A. David & M. W. Reeder (Eds.), *Nations and households in economic growth: Essays in Honor of Moses Abramovitz* (pp. 89–125). New York: Academic.

Ebenstein, A., Harrison, A., & McMillan, M. (2015). *Why are American workers getting poorer? China, trade and offshoring.* NBER Working Paper, No. w21027.

Epstein, R. A. (2008). Happiness and revealed preferences in evolutionary perspective. *Vermont Law Review, 33*(Book 3), 559–83.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness. *The Economic Journal, 114*(467), 641–59.

Geishecker, I. (2008). The impact of international outsourcing on individual employment security: A micro-level analysis. *Labour Economics, 15*(3), 291–314.

Geishecker, I. (2012). Simultaneity bias in the analysis of perceived job insecurity and subjective well-being. *Economics Letters, 116*(3), 319–21.

Geishecker, I., Riedl, M., & Frijters, P. (2012). International outsourcing and job loss fears: An econometric analysis of individual perceptions. *Labour Economics, 19*(5), 738–47.

Hamermesh, D. S. (2001). The changing distribution of job satisfaction. *The Journal of Human Resources, 36*(1), 1–30.

Hanson, G. H., Scheve, K. F., & Slaughter, M. J. (2007). Public finance and individual preferences over globalization strategies. *Economics & Politics, 19*(1), 1–33.

Helliwell, J. F., & Barrington-Leigh, C. P. (2010). Viewpoint: Measuring and understanding subjective well-being. *Canadian Journal of Economics, 43*(3), 729–53.

Horgos, D. (2009). Labor market effects of international outsourcing: How measurement matters. *International Review of Economics and Finance, 18*, 611–23.

Hsieh, C. (2016). Domain importance in subjective well-being measures. *Social Indicators Research, 127*(2), 777–92.

Hummels, D., Munch, J. R., & Xiang, C. (2016a). *Offshoring and labor markets.* NBER Working Paper, No. 22041.

Hummels, D., Munch, J. R., & Xiang, C. (2016b). *No pain, no gain: The effects of exports on effort, injury, and illness.* Working Paper, Purdue University.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–92.

Khun, C., Lahiri, S., & Lim, S. (2015). Do people really support trade restrictions? Cross-country evidence. *Journal of International Trade and Economic Development, 24*(1), 132–46.

Lévy-Garboua, L., & Montmarquette, C. (2004). Reported job satisfaction: What does it mean? *Journal of Socio-Economics, 33*(2), 135–51.

Lü, X., Scheve, K., & Slaughter, M. J. (2012). Inequity aversion and the international distribution of trade protection. *American Journal of Political Science, 56*(3), 638–54.

Muñoz de Bustillo, R., Fernández-Macías, E., Antón, J.-I., & Esteve, F. (2011). *Measuring more than money: The social economics of job quality*. Cheltenham and Northampton: Edward Elgar.

Newport, F. (2012). Most in U.S. know a laid-off worker, but don't fear job loss. Retrieved July 19, 2017 from http://www.gallup.com/poll/154052/know-laid-off-worker-don-fear-job-loss.aspx

Rose, A. K. (2004). Do WTO members have more liberal trade policy? *Journal of International Economics, 63*(2), 209–35.

Scheve, K. F., & Slaughter, M. J. (2001). What determines individual trade-policy preferences? *Journal of International Economics, 54*, 267–92.

Scheve, K. F., Slaughter, M. J. (2004). Economic insecurity and the globalization of production. *American Journal of Political Science, 48*(4), 626–74.

Skalli, A., Theodossiou, I., & Vasileiou, E. (2008). Jobs as Lancaster goods: Facets of job satisfaction and overall job satisfaction. *The Journal of Socio-Economics, 37*(5), 1906–20.

Sousa-Poza, A., & Sousa-Poza, A. A. (2003). Gender differences in job satisfaction in Great Britain, 1991–2000: Permanent or transitory? *Applied Economics Letters, 10*, 691–94.

Sturgeon, T. J. (2013). *Global value chains and economic globalization–Towards a new measurement framework*, A Report to Eurostat. Boston: MIT.

UNCTAD (2004). *World investment report 2004: The shift towards services*. New York, NY: UN Conference on Trade and Development.

UNCTAD (2015). *Key statistics and trends in international trade 2014*. New York and Geneva: UN.

Veenhoven, R., & Hagerty, M. (2006). Rising happiness in nations 1946–2004: A reply to Easterlin. *Social Indicators Research, 79*(3), 421–36.

WTO (2015). *International trade statistics 2015*. Geneva: WTO.

Yeh, H.-J. (2015). Job demands, job resources, and job satisfaction in East Asia. *Social Indicators Research, 121*(1), 47–60.

Gender Pay Gaps in Domestic and Foreign-Owned Firms

Iga Magda and Katarzyna Sałach

1 Introduction

A number of studies have shown that, contrary to most theoretical predictions, the gender wage gap tends to be larger in foreign-owned companies than in domestically owned firms. This larger gender wage gap is found both when the raw differences in the average wages of men and women

This paper benefited from the financial support of the National Science Center, Poland (grant number 2013/10/E/HS4/00445). We use the data provided by Statistics Poland. Statistics Poland has no responsibility for the results and conclusions, which are those of the authors.

I. Magda (✉)
SGH Warsaw School of Economics, IBS Institute for Structural Research, Warsaw, Poland
e-mail: iga.magda@sgh.waw.pl

K. Sałach
University of Warsaw, Warsaw, Poland
e-mail: ksalach@wne.uw.edu.pl

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_8

are measured, and when the pay gap is adjusted by taking into account differences in observable individual, job, and firm characteristics. So far, however, neither theoretical nor empirical research has provided a convincing explanation for why women are more disadvantaged in terms of pay if the company they work for is owned by foreign investors.

Our study has two main goals. First, using different methodological approaches to calculating gender pay gaps, we aim to determine whether these gaps are indeed larger in foreign-owned firms than in domestically owned firms, and whether this pattern holds both for low- and high-wage earners. Second, we seek to shed light on the factors that could explain the differences in the size of the gender pay gap depending on firm ownership. In particular, we are interested in learning whether firm-level wage policies differ between domestically owned and foreign-owned firms.

Our study focuses on Poland, which, like other Central and East European countries, benefited from large foreign direct investment (FDI) inflows after the economic transition in the early 1990s. As a consequence, foreign-owned firms are well established in the Polish market: 16% of entities employing 10 or more employees in Poland in 2014 were at least partially owned by foreign investors (CSO, 2015).[1] These entities employed 14% of all paid employees in Poland and offered them wages that were, on average, 60–70% higher than those offered by domestically owned firms. These features make examining the gender pay gap in foreign-owned firms particularly interesting.

Four main findings emerge from our study. First, gender wage gaps are significantly larger in the foreign sector than in the domestic sector. The raw gender wage gap is more than 2.7 larger in foreign-owned than in domestically owned companies. Adjusting for firm and worker characteristics and for firm fixed effects decreases this difference considerably, although the adjusted gender wage gap is still 45–80% larger in the foreign-owned than in the domestically owned sector. Moreover, the gender wage gaps are larger in foreign-owned firms than in domestically owned companies at every quantile of the wage distribution, although the differences are the greatest at the two ends of the wage distributions and are the smallest in the middle. Second, while in the foreign-owned sector the returns to individual, job, and firm characteristics earned by women

[1] The largest shares of foreign capital were invested in the manufacturing sector, the wholesale and retail trade, motor vehicle and motorcycle repair, and information and communication.

are much lower than the returns earned by men, foreign-owned firms appear to pay higher firm-specific wage premia to women than to men, thereby narrowing within-firm gender wage inequality. These patterns differ from those observed in the domestic sector, where firm wage premia widen within-firm wage distributions and contribute to overall wage inequality. However, despite this divergence in firm and price effects in the two ownership sectors, the median and the average levels of within-firm wage inequality are still higher in foreign-owned than in domestically owned firms. These findings hold for both for low- and high-paid workers in the two ownership sectors. Third, the estimates of gender wage inequality can be susceptible to the reference group chosen. In particular, we argue that the greater heterogeneity in worker and firm characteristics in the domestic sector drives the difference in the estimates of the gender wage gap based on male versus female characteristics. Fourth, we find that the two ownership sectors display similar patterns of gender pay differences among firms with large within-firm wage gaps, but different patterns of gender pay differences at the lower tail of the within-firm wage gap distribution. Specifically, we show that there is a larger share of domestically owned firms in which men earn less than "similar" women.

This chapter is organised as follows. In Sect. 2, we review the relevant literature. In Sect. 3, we present the data we use. The methodological approach and the results are discussed in Sect. 4. Section 5 concludes.

2 Firm Ownership and Gender Wage Differentials

Economic theory suggests that gender pay gaps should be smaller among foreign-owned companies than among domestically owned firms. This prediction is based on the observation that, compared to their domestic counterparts, foreign-owned firms are more likely to operate under highly competitive market conditions. As such, in line with the predictions of the personal taste hypothesis, discrimination is expected to be (more) costly for foreign-owned firms (Becker, 1957; Arrow, 1973). These theoretical arguments are further reinforced by the assumption that the weaker product market competition that is enjoyed by domestic companies, and by publicly owned firms in particular, could create opportunities for higher rents, which may be shared with employees. To the extent that these domestic firms prefer to employ men and to reward them more than women (gender differences in rent-sharing have been confirmed by Nekby, 2003), gender wage gaps should be larger in domestic firms than in

foreign-owned companies. Apart from these competition theory considerations, the expectation that gender pay gaps would be smaller in foreign-owned firms is supported by trade theory, which posits that the ability of foreign-owned firms to engage in gender-based pay discrimination is reduced (Black & Brainerd, 2004). Again, this expectation is based on the observation that foreign-owned firms tend to import and export products, whereas domestic companies tend to be oriented towards the domestic market, where competition is weaker.

The empirical evidence regarding these assumptions is inconclusive: the theoretical link between the (higher) degree of market competition and the (smaller) size of the gender labour market gap has been confirmed by Black and Strahan (2001); Meng (2004); Zweimueller et al. (2008); and Lovasz (2008). By contrast, Li and Dong (2011) found that firms that have larger gender wage premia are more likely to operate in industries subject to fierce competition.

There could be other reasons why the gender pay gap might be smaller in foreign-owned firms. First, these companies may have firm-level policies regarding childbearing and childcare that result in smaller gender wage differentials. Family-friendly practices in the workplace can help to close the gender pay gap (Felfe, 2012). Foreign-owned firms may be more likely than domestically owned companies to support both equal pay and equal promotion legislation and family-friendly workplace solutions (Kodama et al., 2018). If highly educated women are selected into foreign-owned companies because these firms have a flexible approach to work-life balance,[2] the pay gaps in these companies should be smaller. We assume that these transmission mechanisms are particularly important for Poland, as approximately 90% of all foreign capital that has been invested in Poland came from EU countries (with the biggest shares coming from the Netherlands (18%), Germany (16%), and France (15%) (Statistics Poland, 2017)). We would expect to find that these firms "import" their pay

[2] To the best of our knowledge, there are no studies that have investigated how work-life balance policies differ between domestically and foreign-owned firms in Poland or other Central Eastern European countries. However, when we refer to data from the 2018 Labour Force Survey ad hoc module on reconciliation between work and family life for Poland, we find that one of the dimensions of work-life balance—namely, working time flexibility—is more than twice as common among private-sector workers as among public-sector workers. While there is no domestic/foreign dimension, we can see that the types of employees over-represented in foreign-owned firms (younger, better educated, working in information or communication) enjoy more working time flexibility.

policies from their home countries, which tend to have much smaller adjusted gender pay gaps than Poland (Christofides et al., 2013).

However, some authors have argued that gender pay gaps are likely to be larger in foreign-owned companies because these firms often require employees to work long hours. Such demands tend to benefit men, who are more likely than women to be willing to work long hours and to adjust to a flexible schedule (Goldin, 2014; Vahter & Masso, 2019). Similarly, Bøler et al. (2018) have suggested that exporting firms may require their workforce to have a greater degree of employer-centred working time flexibility, as employees may need to work with customers in different time zones. Since exporting firms are more likely to be foreign owned than domestically owned, such demands on workers might also contribute to the observed differences in the gender pay gaps of domestically owned and foreign-owned workplaces.

The empirical literature that refers explicitly to differences in the gender pay gaps in domestically owned and foreign-owned firms is limited. It is widely acknowledged that foreign firms usually offer wage premia (attributed to the technology, capital, and competition externalities of multinational companies) that have a direct impact on the foreign-domestic pay gap (Hijzen et al., 2013). However, it is less obvious whether (and, if so, why) these foreign-ownership wage premia are higher or lower for men than for women; and, thus, whether the gender pay gap is increased or decreased by FDI inflows and ownership structure. Most of the previous research that addressed these questions investigated conditions in China from a microeconomic perspective. These studies found that wage premia are indeed higher for men than for women in the foreign ownership sector, and that the gender pay gaps are therefore larger in foreign-owned firms than in domestically owned companies (Liu et al., 2000; Maurer-Fazio & Hughes, 2002; Rickne, 2012). Braunstein and Brenner (2007) found that while FDI inflows benefited the wages of women more than those of men in the mid-1990s, this pattern reversed in the early 2000s. It is worth emphasising that the mechanisms that operate in a developing country might not be present in a more advanced context, where FDI inflows may not translate into more women entering the labour market or attaining higher levels of education (Seguino & Grown, 2006; Oostendorp, 2009). To the best of our knowledge, there is little evidence regarding this relationship for European countries. The main exceptions are Zulfiu-Alili (2014), who found that gender wage gaps are larger in foreign-owned firms than in domestically owned companies in Macedonia,

and Vahter and Masso (2019), who observed a similar pattern in Estonia. Earle et al. (2018) found that the foreign acquisition of Hungarian firms had a similar impact on the wages of men and women. Thus, it appears that foreign ownership is not necessarily associated with a smaller or a bigger gender pay gap (at least in firms that are taken over).

We add to the studies on the association between gender pay gaps and firm ownership by linking our research to the literature on within-firm wage inequality and its gender dimension. The existing studies differ in terms of both their methodological approaches and the results they provide, although most find that women earn only a portion of the firm wage premia earned by men. Meng (2004) and Meng and Meurs (2004) decomposed the gender wage differentials into standard endowment and coefficients effects and additional firm wage premia to show that firm wage policies reduced unexplained gender wage gaps to a much greater extent in Australia than in France. Heinze and Wolf (2010) were able to partly capture the non-random selection of women into firms by calculating the gaps separately for each firm in their study. However, they found only a small difference between the mean cross-section gender wage gap within firms and the mean overall gender wage gap. By contrast, Magda and Cukrowska-Torzewska (2019) found that in Poland, both the raw and the adjusted gender gaps were much smaller when men and women were compared within firms rather than across firms.[3] In their recent influential paper, Card et al. (2016) estimated employer and employee wage fixed effects (firm wage premia) and found that both sorting across firms (i.e. women's higher probability of working at firms that pay low wages) and differences in within-firm bargaining (i.e. women receiving less of the wage premium than men) contributed to the gender wage gap. In a similar vein, Javdani (2015) separated the effects of female segregation into low-paying firms and female segregation into low-paying jobs within firms and concluded that the former effects were a major driver of gender wage inequality. Other studies of within-firm gender wage gaps include Song et al. (2019), Bayard et al. (2003), Reilly and Wirjanto (1999). We contribute to this literature with empirical findings showing the need to account for firm-level

[3] Other studies on gender pay gaps in Poland neither focused on the firm-level wage premia, nor distinguished between foreign and domestic firms, which are our major contributions. The recent studies on this topic looked at the measurement of the adjusted pay gap (Matysiak et al., 2010; Goraus et al., 2017), its changes over time (Tyrowicz & Smyk, 2019), and the role of occupational segregation (Strawiński et al., 2018).

characteristics, as these characteristics affect the differences in the estimates of economy-wide and within-firm gender wage gaps. We also show that firms' heterogeneity—in particular along the domestic/foreign ownership divide—leads to the sensitivity of the results to the choice of reference group in the decompositions of gender wage gaps.

3 Data and Descriptive Statistics

We use data from the Structure of Wages and Salaries by Occupations (SWSO) survey conducted by Statistics Poland in 2014.[4] The SWSO is a large, linked employer-employee dataset that covers organisations employing ten or more employees and that provides information on both the yearly and the monthly (during the reference month of October) earnings of individuals. The dataset also contains information on the number of normal and overtime hours employees have worked, and on a range of individual characteristics (e.g. gender, age, education, occupation, experience, tenure) and firm characteristics (e.g. NACE, type of ownership [public/private and domestic/foreign ownership], coverage by collective pay agreement, and firm size). Because we are interested in comparing the gender wage gaps in domestically owned and foreign-owned firms in the private sector, we restrict our sample to companies with one ownership type only (i.e. we exclude firms with mixed ownership).[5] For the same reason, we exclude public sector firms entirely. The final sample includes 343,143 individual observations: 222,203 in the domestic sector and 120,940 in the foreign sector. We use sample weights that reflect the survey's two-stage sampling procedure (at the firm level and at the worker level). We calculate gender wage gaps using data on hourly wages, which we compute as the yearly salary divided by the number of hours worked yearly. We include in the salary any compensation from overtime, awards, additional fees, and statutory bonuses.

Foreign-owned firms account for 14.3% of all of the firms in our data, and employ 30% of all of the workers in the sample. Clearly, there are differences in the structure of the workforce depending on ownership type (Table 1). Women constitute a minority of the workforce in both the

[4] The survey is conducted every two years. The analyses for the years 2008, 2010, and 2012 yielded similar results, and are available from the authors upon request.

[5] We also exclude observations with wages that we consider mistakenly low (far below the minimum wage) and those that claim that the unit size is less than ten workers (they should not be captured by the survey by design).

Table 1 Descriptive statistics of selected variables, 2014

	Domestic	*Foreign*
Female	0.40	0.43
	(0.49)	(0.49)
Age	40.3	36.7
	(11.3)	(9.94)
Primary education	0.07	0.07
	(0.25)	(0.26)
Basic vocational education	0.30	0.18
	(0.46)	(0.39)
Secondary education	0.38	0.36
	(0.49)	(0.48)
Tertiary education	0.24	0.39
	(0.43)	(0.49)
Job experience	15.7	12.6
	(11.6)	(10.1)
Tenure	7.9	7.1
	(8.4)	(7.4)
Firm size (average number of workers)	334	1136
	(651)	(2979)
Fixed-term contracts	0.39	0.28
	(0.49)	(0.45)
Collective agreements (both firm-level and industry-level)	0.38	0.34
	(0.48)	(0.47)
Men, average hourly wage (PLN)	19.77	34.80
	(18.9)	(36.1)
Women, average hourly wage (PLN)	17.39	25.59
	(12.6)	(24.4)
Raw gender wage gap (difference in wages of men and women as % of men's wages)	12.0%	26.5%
Number of firms	6226	1269
Number of individual observations	222,203	120,940

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: Standard deviations in parentheses. Sample weights used

foreign-owned and the domestically owned companies, though their share is slightly larger in the foreign-owned companies. Employees of the foreign firms are, on average, three years younger and better educated than those of the domestic firms. Compared to their counterparts who work for domestic firms, employees who work for foreign establishments are less likely to be employed on fixed-term contracts, and they are more likely to work for a large organisation.

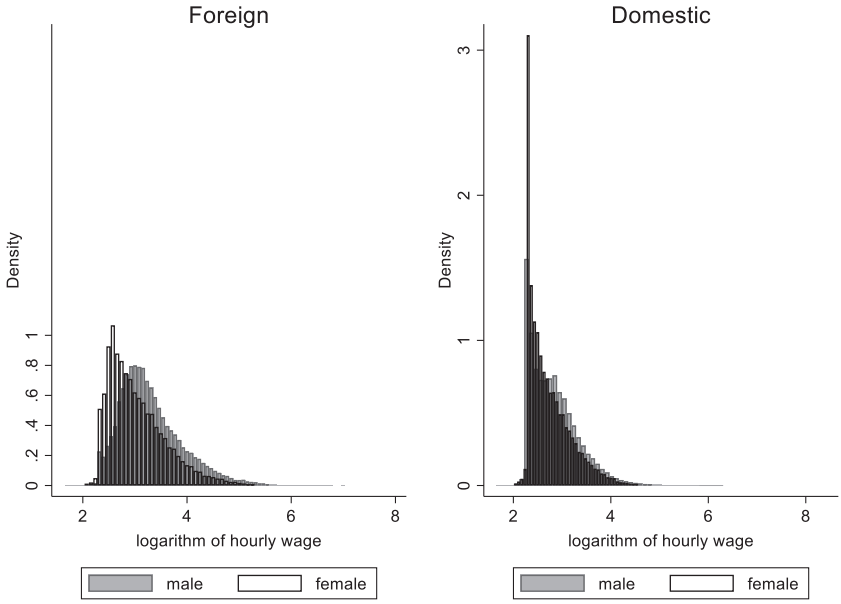

Fig. 1 Men's and women's distribution of log wages in foreign- and domestically owned firms. (Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data)

In terms of wages, men who work for foreign-owned companies earn on average 76% more than men who work for domestically owned firms. Among women, the corresponding difference is 47%. Furthermore, in both sectors, the distribution of female wages is shifted to the left of the male distribution, but this shift is greater in the foreign sector (Fig. 1). Thus, gender wage inequality appears to be considerably higher in foreign-owned than in domestically owned companies.

4 Gender Wage Gap and Firm Ownership

We start the analysis of gender wage gaps in domestically and foreign-owned firms by comparing raw gaps and gaps adjusted for individual and workplace characteristics using classical Mincerian regressions (Section "Adjusted Gender Wage Gaps and Firm Ownership"). Next, we attempt to shed more light on the potential sources of variation in

the size of the gender wage gaps in the two ownership sectors (Section "Decomposition of the Gender Wage Gap"), and focus more on the role firms play (Section "Firm-Specific Effects and Gender Wage Gaps"). Finally, we analyse the gender pay gaps among low- and high-paid workers using quantile regressions (Section "Unconditional Quantile Regression").

Adjusted Gender Wage Gaps and Firm Ownership

In the first step, we calculate the raw gender wage gaps; that is, the difference in the average hourly wages of men and women, expressed as the percentage of men's wages. We do so separately for the two types of firm ownership: domestic and foreign. The raw gender wage gap in Poland is considerably larger among individuals who are working in foreign-owned than in domestically owned firms (Fig. 1, Table 1). Although women have lower wages than men in both sectors, the raw gender wage gap is more than twice as large in the foreign-owned firms as it is in the domestically owned firms (26.5% and 12.0%, respectively, Table 1).

Obviously, the raw gender wage gap is not the most suitable measure of gender wage inequality. While the size of the gender pay gap varies substantially between the two ownership sectors, this pattern may be explained in part by differences in the composition of male and female workers in domestically and foreign-owned firms. To eliminate this effect, we calculate the gender wage gap adjusted for the characteristics of workers, jobs, and firms. We use a traditional Mincer wage regression with the logarithm of the hourly wage as a dependent variable. We include in the models an interaction term between gender (female) and type of ownership (foreign), and we estimate them using weighted least squares (WLS).[6] As a robustness check, results of regressions estimated separately for domestic and foreign firms are presented in the Appendix, Tables 5, 7, and 8

Our models contain a set of standard control variables, which we group into four categories: individual-level characteristics (age, education, experience), job-level characteristics (occupation, type of job contract, tenure, and part-time/full-time position), firm-level characteristics (firm size, NACE sector, collective bargaining coverage), and co-worker characteristics (share of women within a given firm, share of workers aged 55 or older, share of workers aged 34 or younger, share of workers with tertiary

[6] Using sample weights, as described in Sect. 3.

education). Model 1 contains individual-level, job-level, and firm-level characteristics. Model 2 adds to these variables co-worker characteristics. Model 3 uses firm fixed effects instead of firm-level and co-worker characteristics. Finally, we run model 3 on the common support, as explained below (results reported as model 4). In models 1 and 2, we cluster standard errors at the firm level.

In order to address the possibly limited overlap in the observable characteristics of male and female employees, we calculate the common support. In doing so, we follow Lechner and Strittmatter (2019): we estimate the propensity score—the conditional probability of being a female given the value of covariates—and then drop observations for both genders when the value of the propensity score is higher than the 99th percentile of the propensity score distribution in the male subpopulation. In our case, this procedure resulted in 2.8% of the sample being dropped.

We must admit, that—in line with many analyses of this type—we do not observe many factors that could be correlated with both gender and wages, such as personality traits, the propensity to negotiate wages, gender role attitudes, or absenteeism. We implicitly assume in our analysis that gender is independent of these unobservables conditional on observables, which is disputable. Moreover, following a common assumption in the gender wage gap literature, we treat job experience and occupation as exogenous, which may lead to inconsistent parameter estimates (Kunze, 2008). This is a limitation of our study, and we acknowledge it.

Table 2 presents a summary of the results, and detailed estimates are presented in Appendix, Table 6. First, as a benchmark, we calculate the raw difference in the logarithms of hourly wages, and find that it amounts to 0.100 in the domestic sector and 0.273 in the foreign sector. Model 1, which includes individual-level, job-level, and firm-level characteristics as covariates, yields an adjusted gender wage gap equal to 0.164 in the domestic sector and to 0.249 in the foreign sector. Model 2 also includes a set of co-worker characteristics in the covariates list. It turns out that these particular variables play an important role in explaining gender wage gaps, as in model 2, these gaps decrease to 0.107 in domestically owned firms and to 0.194 in foreign-owned firms. Thus, it appears that taking co-worker characteristics into account solves the omitted variables problem to a certain extent and captures a portion of firms' unobserved heterogeneity. Indeed, the estimated adjusted gender wage gaps are found to be similar once we run a model with firm fixed effects (model 3): 0.116 in the domestic sector and 0.172 in the foreign sector. Finally, in model 3

Table 2 WLS-adjusted gender wage gaps in domestically and foreign-owned firms

Ownership	*Raw difference in logarithms*	*Model 1*	*Model 2*	*Model 3*	*Model 4*
Domestic	0.100	0.164	0.107	0.116	0.117
Foreign	0.273	0.249	0.194	0.172	0.170
Model specification					
Individual characteristics	–	Yes	Yes	Yes	Yes
Job characteristics	–	Yes	Yes	Yes	Yes
Firm characteristics	–	Yes	Yes	No	No
Co-worker characteristics	–	No	Yes	No	No
Firm fixed effects	–	No	No	Yes	Yes
Common support	No	No	No	No	Yes

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: The full set of estimates is available in Appendix, Table 6. The models are estimated on a pooled dataset with both domestically and foreign-owned firms, and use the female#foreign interaction term. The regression results estimated separately for domestic and foreign firms as a robustness check are presented in Appendix, Tables 7 and 8

re-estimated on the common support sample (model 4), the results are very similar. It must be noted, however, that in practice, the wage gap estimates from models 1–4 cannot be interpreted as either wage gains for men or wage losses for women. If there are more or less equal numbers of advantaged and disadvantaged individuals in the sample—as in our case—the estimates are closest to the average wage gap (Słoczyński, 2019). We will attempt to shed more light on these estimates in Section "Decomposition of the Gender Wage Gap".

Thus, we find that the adjusted gender wage gaps are indeed larger in the foreign-owned companies, but that the difference between the two ownership sectors is much smaller than the raw gaps would suggest. Moreover, distinct patterns emerge in the two ownership sectors. First, once we adjust for worker and firm characteristics, the gender pay gap increases in the domestic sector (compared to the raw gender pay gap), but decreases in the foreign sector. Second, firm-level unobserved heterogeneity (which is captured by co-worker characteristics to a similar extent as by firm fixed effects) plays an important role in shaping gender wage inequality. However, it appears that firm-level policies have different outcomes in domestically and foreign-owned firms. To gain a better understanding of the nature of the gaps in the two sectors and the role firms

play, we decompose gender wage gaps in Section "Decomposition of the Gender Wage Gap" and analyse within-firm gender wage gaps in Section "Firm-Specific Effects and Gender Wage Gaps".

Decomposition of the Gender Wage Gap

We draw on the standard generalised linear decomposition of the difference in wages between men and women (Oaxaca & Ransom, 1994). The difference (the wage gap) is decomposed into the composition effect ("the explained part") that arises due to differences in the characteristics of men and women and the coefficient effect ("the unexplained part," or, in this setting, the adjusted gender wage gap) that arises due to different rewards from the characteristics. It takes the form of:

$$\begin{aligned}\mathrm{E}(y|f=0)-\mathrm{E}(y|f=1)=&\left[\mathrm{E}(X|f=0)-\mathrm{E}(X|f=1)\right]\beta^{*}+\mathrm{E}(X|f=0)\left(\beta_{m}-\beta^{*}\right)\\&+\mathrm{E}(X|f=1)\left(\beta^{*}-\beta_{f}\right),\end{aligned}$$

$$\begin{aligned}\mathrm{E}(y|f=0)-\mathrm{E}(y|f=1)=&\left[\mathrm{E}(X|f=0)-\mathrm{E}(X|f=1)\right]\beta^{*}\\&+\mathrm{E}(X|f=0)\left(\beta_{m}-\beta^{*}\right)+\mathrm{E}(X|f=1)\left(\beta^{*}-\beta_{f}\right),\end{aligned}\quad(1)$$

where by y we denote wage (in logarithm); X is the vector of all individual and workplace characteristics; f is an indicator for female workers; β_m and β_f are the vectors of regression coefficients for men and women, respectively; and β^* is defined as a weighted average of the coefficient vectors:

$$\beta^{*}=\Omega\beta_{m}+\left(I-\Omega\right)\beta_{f}.\quad(2)$$

In the latter equation, Ω is a weighting matrix and I is an identity matrix. The original decomposition equations proposed in seminal papers by Blinder (1973) and Oaxaca (1973) represent special cases of the generalised equation in which Ω is either a null matrix or an identity matrix, respectively.

However, the choice of Ω is not innocuous. Słoczyński (2019) and Goraus et al. (2017) summarised the possible options. We stick to the two the most popular ones: namely, $\Omega = 1$, where male coefficients are taken as a reference (and we thus estimate the average wage loss for women); and $\Omega = 0$, where female coefficients are taken as a reference (estimating the

Table 3 Gender wage gaps in domestically and foreign-owned firms, adjusted for firm and worker characteristics: Oaxaca-Blinder decomposition results

	Total predicted gap	*Ω = 0*		*Ω = 1*	
		Unexplained (average wage gain for men)	*Explained*	*Unexplained (average wage loss for women)*	*Explained*
Domestic	0.100	0.129	-0.029	0.088	0.012
		(0.003)	(0.003)	(0.003)	(0.003)
Foreign	0.273	0.183	0.090	0.169	0.104
		(0.003)	(0.005)	(0.003)	(0.004)

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: The covariates for the decomposition are the same as those included in model 2. Bootstrap standard errors in parentheses

average wage gain for men). Other possible choices of Ω include weighting each coefficient by the proportion of the same gender (Cotton, 1988); Ω = 0.5, or the simple average of both (Reimers, 1983); weighting each coefficient by the proportion of the opposite gender (Słoczyński, 2019); and taking the coefficients from a pooled regression, with a gender dummy (Fortin, 2008) or without it (Neumark, 1988).

The decomposition is based on model 2.[7] The results confirm that while the adjusted (unexplained) gender pay gap is greater in the foreign-owned firms than it is in the domestically owned firms, this ownership gap is much smaller than the raw gender wage gaps would suggest (Table 3). At the same time, the decomposition reveals another difference between the two ownership types: namely in the foreign-owned firms, the average wage gain for men is similar to the average wage loss for women (0.183 and 0.169 account for 67% and 62% of the total gap, respectively); whereas in the domestically owned firms, the average wage gain for men exceeds 129%, and the explained part of the gap is negative. This means that while female workers have "better" characteristics than males in the domestic sector, and this advantage should narrow the gap, the returns to these characteristics paid to men are so much higher than the returns paid to women that they more than counterbalance the better endowments of women. The picture changes once we focus on the counterfactual wages of women, which would be observed if they were paid according to the

[7] We additionally run the decomposition on common support only. The results are robust, and are available from the authors upon request.

wage structure of men (average wage loss for women). The unexplained gender pay gap is smaller.

Altogether, the results confirm that the unexplained ("adjusted") gender pay gap is larger in the foreign sector than in the domestic sector. However, the difference is much smaller than the raw gender pay gaps would suggest, as the different characteristics of men and women employed in foreign-owned companies explain part of the gap. From a methodological perspective, our findings show that even if the proportions of the advantaged and disadvantaged groups (men and women) are similar in the domestic and the foreign sector, the choice of the reference group matters more in domestically owned firms.[8]

Firm-Specific Effects and Gender Wage Gaps

In the next step, we want to shed more light on the effects of firms on gender wage inequality. To this end, we use two strategies of looking at within-firm gender pay gaps. First, taking advantage of our linked employer-employee data, we calculate the gender pay gaps separately for each firm using WLS regressions with a set of covariates that include personal and job characteristics, as in model 1. To do so, we had to exclude the smallest firms in order to have a reasonable number of observations for the calculation. We therefore excluded firms with fewer than 100 workers and fewer than 10 observations of men or women, which resulted in the number of observations per firm ranging from 40 to 868, with the median at 100. The distribution of the estimated coefficients associated with the male dummy (which are interpreted as gender pay gaps) in each firm is plotted in Fig. 2, separately for each sector of ownership.

On average, within-firm gender wage gaps are slightly larger in foreign-owned firms. The difference between the average size of the within-firm gender wage gap in the two ownership sectors is small, but statistically significant (0.12 versus 0.10 in foreign- and domestically owned firms, respectively). The analysis also reveals that the within-firm gender wage

[8] The results are similar when we run the model on common support; thus, even when the overlapping support assumption is ensured. Still, the combination of observed workers and workplace characteristics is much different for men and women in domestically owned firms than it is for their counterparts in foreign-owned firms. We calculated a simple Duncan dissimilarity index (Duncan & Duncan, 1955). It provided evidence of a significantly higher degree of worker dissimilarity by gender in the domestic sector than in the foreign sector. This observation holds regardless of the combination of individual, job, and firm characteristics we take into account.

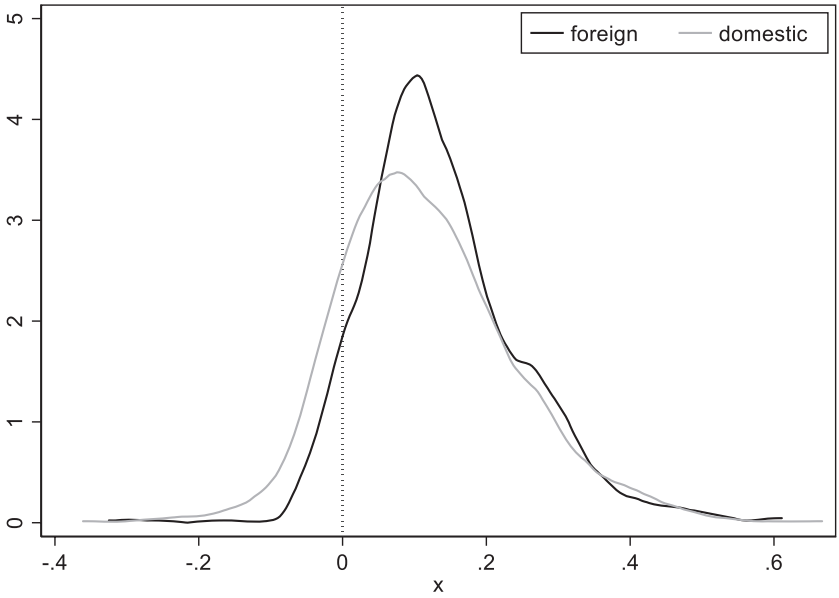

Fig. 2 Kernel density of within-firm gender wage gap estimates. (Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data. Note: Firms employing at least 100 workers and with at least 10 observations of each gender in the sample. A positive gap means that men earn more than "similar" women)

gaps are more dispersed among the domestically owned than the foreign-owned companies. The standard deviation of the gender wage gap is larger in the domestic sector (0.12) than in the foreign sector (0.11), and the difference is again statistically significant. Interestingly, 16% of domestically owned firms have negative gender pay gaps, meaning that in these firms, men earn less than "similar" women. This is the case for 7% of the foreign-owned companies.[9] Thus, we find that among medium- and

[9] The firms that advantage women in terms of pay do not seem to be particularly different from the other firms, both domestically and foreign-owned, in terms of observable characteristics such as industry, mean age of workers, workers' mean level of education, or workers' experience. There are two exceptions: both domestically and foreign-owned firms that pay particularly high wage premia to women tend to be smaller and to have lower average wages.

large-sized firms, the gender pay gaps are, on average, quite similar in domestically and foreign-owned firms, although the domestically owned sector is more heterogeneous.

In the second step, we draw on Meng and Meurs (2004) and extend the Oaxaca-Blinder decomposition to isolate firm effect from the unexplained part, so that *total predicted gap = endowment effect (explained) + price effect (unexplained) + firm effect:*

$$\begin{aligned} \mathrm{E}(y|f=0)-\mathrm{E}(y|f=1) = & \left[\mathrm{E}(X_c|f=0)-\mathrm{E}(X_c|f=1)\right]\beta^* \\ & +\left[\mathrm{E}(X_{firm}|f=0)-\mathrm{E}(X_{firm}|f=1)\right]\beta^* \\ & +\mathrm{E}(X_c|f=0)(\beta_m-\beta^*)+\mathrm{E}(X_c|f=1)(\beta^*-\beta_f) \\ & +\mathrm{E}(X_{firm}|f=0)(\beta_m-\beta^*)+\mathrm{E}(X_{firm}|f=1)(\beta^*-\beta_f) \end{aligned} \quad (3)$$

where all symbols denote the same as in Eq. (1), with the exception of the vector of characteristics *X,* which is now split into firm dummies (X_{firm}) and other controls (X_c); that is, individual, job, co-worker, and firm characteristics (except for the firm indicator). The firm effect, which is the main component of our study, is given by $\mathrm{E}(X_{firm}|f=0)(\beta_m - \beta^*) + \mathrm{E}(X_{firm}|f=1)(\beta^* - \beta_f)$ part of Eq. (3). Please note that the explained firm component $[\mathrm{E}(X_{firm}|f=0) - \mathrm{E}(X_{firm}|f=1)]\beta^*$ (stemming from the fact that women and men work in different firms) contributes to the endowment effect.

We use a fixed effects model (model 3) to estimate the firm-specific wage premia, for men and women separately. Table 4 summarises the results again separately for wage gains for men ($\Omega = 0$) and wage losses for women ($\Omega = 1$). Compared to the Oaxaca-Blinder decomposition from the previous section (in which we did not include firm fixed effects), the unexplained component is now accounted for by the price effect (reflecting different returns to individual characteristics among men and women) and the firm effect, which captures differences in the wage premia paid by particular firms to men and women in these firms. We find that in the foreign-owned sector, the returns to individual, job, and firm characteristics that women earn are much lower than the returns earned by men. This is reflected in the price effect, which amounts to 189–200% of the total difference in the average wages of men and women in the foreign-owned companies. In contrast, the firm effect is also large in foreign-owned companies (approximately 130%) but is negative—which means that foreign-owned companies tend to have narrower within-firm gender wage gaps.

Table 4 Gender wage gaps in domestically and foreign-owned firms, adjusted for firm and worker characteristics: decomposition into endowment, price, and firm effects

	Domestic		*Foreign*	
	Value	*%*	*Value*	*%*
Ω = 1				
Total difference	0.107		0.274	
	(0.002)		(0.003)	
Price effect	-0.040	-37	0.517	189
	(0.000)		(0.000)	
Endowment effect	0.005	5	0.092	34
	(0.002)		(0.004)	
Firm effect	0.141	132	-0.335	-122
	(0.001)		(0.001)	
Ω = 0				
Total difference	0.107		0.274	
	(0.002)		(0.004)	
Price effect	-0.021	-20	0.547	200
	(0.000)		(0.000)	
Endowment effect	-0.020	-19	0.100	37
	(0.002)		(0.004)	
Firm effect	0.148	139	-0.373	-136
	(0.001)		(0.001)	

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Note: Decomposition method based on Meng and Meurs (2004). Bootstrap standard errors in parentheses. We discarded firms with a 0% or a 100% share of female employees in order to be able to calculate firm effects

The opposite is the case for the domestically owned firms, in which the within-firm gender wage gaps are wider. This is reflected in the positive firm effect, which amounts to approximately 135% of the average gender wage gap in domestically owned companies.

To sum up, the gender wage gap is much larger among workers in foreign-owned companies than it is among workers in domestically owned firms. In the foreign-owned firms, the within-firm gender wage differentials are narrower, although the firm-specific wage premia they pay to women are not sufficient to compensate for the gender wage inequality observed in the entire sector. Domestically owned firms offer higher firm-specific wage premia to men than to women, which increases within-firm wage inequality and contributes to the overall gender wage gap in the domestic sector. How can we reconcile the findings of positive firm effects and negative price

effects among foreign-owned firms? In other words, why do female workers working in foreign-owned firms earn so much less than their male colleagues if these companies have narrower within-firm gender wage gaps? Apart from the omitted variable bias, the sorting channel is one of the potential explanations: that is, women are likely to be underrepresented at firms that pay higher firm-specific wage premia (Card et al., 2016).[10] The difference in the size of the endowment effect observed in the domestically and foreign-owned firms lends support to this hypothesis.

Unconditional Quantile Regression

Finally, we look at how the gender pay gap changes along the wage distribution, in the domestically and foreign-owned firms separately. To this end, we use unconditional quantile regression (UQR), as proposed by Firpo et al. (2009). Unlike conditional quantile regression (CQR), UQR defines quantiles of the variable of interest (logarithm of wages in our case) *prior to* regression. This means that the inclusion of any covariates in the regression helps to net out their effect on the relationship of interest (i.e. in our case, between wages and gender), but it has no effect on which observations are assigned to given quantiles of the wage distribution (Killewald & Bearak, 2014).

The UQR method consists of transforming the dependent variable into the re-centred influence function (RIF) of the unconditional quantile of the wage distribution, and then regressing it on the list of explanatory variables. Standard OLS regression can be used. The RIF takes the form (Firpo et al., 2009):

$$RIF\left(Y;,q_\tau;,F_Y\right) = q_\tau + \frac{\tau - \mathbf{1}_{\{Y \le q_\tau\}}}{f_Y\left(q_\tau\right)}$$

[10] While the formal test of this pattern is beyond the scope of our paper, we run simple OLS estimates of (1) the link between the male wage premium, the share of women in the firm, and firm ownership; and (2) the probability of being employed at a high-/low-paying firm (defined by above the 75th/below the 25th percentile of the male wage premium distribution), controlling for individual- and firm-level characteristics, and interacting female with the foreign/domestic dummy. We find that Polish female workers are less likely to be employed at high-paying foreign-owned companies, but not at high-paying domestically owned companies. This might suggest that the sorting channel is more relevant in foreign-owned companies. We must note, however, that these results are based on correlation, and not causality. The data we use do not have a longitudinal dimension that would allow us to study job switching, as Card et al. do. The results are available from the authors upon request.

where τ is a given quantile, q_τ is the value of the outcome variable, Y, at the τ th quantile, $f_Y(q_\tau)$ is the density of Y at q_τ, F_Y is the cumulative distribution function of Y and $\mathbf{1}_{\{Y \leq q_\tau\}}$ is a dummy variable indicating whether the value of the outcome variable is below q_τ. We bootstrap the standard errors to diminish the uncertainty involved in the estimation of the RIF (Killewald & Bearak, 2014). We apply the UQR to our model 3.

The gender wage gaps are larger in foreign-owned firms than in domestically owned companies in every quantile of the wage distributions (Fig. 3 plots the results). However, the differences are greatest at the two ends of

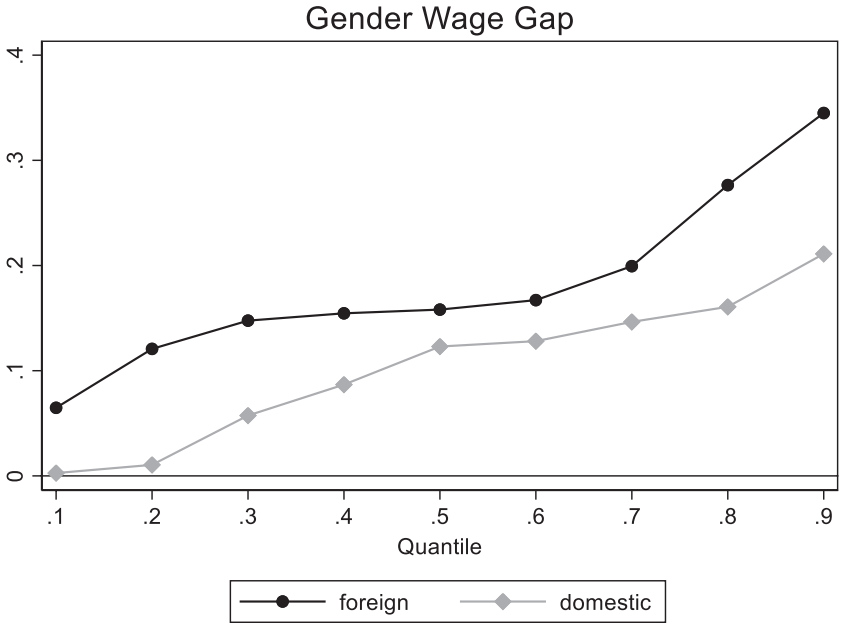

Fig. 3 Gender pay gaps by quantiles, in foreign-owned and domestically owned firms. (Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data. Note: The quantiles are defined prior to regression, and separately for domestically and foreign-owned firms; thus, the absolute wage levels in each quantile are different between the two sectors (and higher in the foreign sector). For clarity, the absolute value of the gender pay gaps is plotted (with the notation we use throughout the paper, they are negative))

the wage distributions and are smallest in the middle. There is virtually no pay gap among low-paid men and women working in the domestic sector, since they are paid the minimum wage. Wages are higher in the foreign-owned companies, and the minimum wage is binding only for a small fraction of workers in these firms.[11] In the first decile of wages in the foreign-owned companies, the adjusted pay gap among men and women does not exceed 7%. The pay gaps increase along the wage distribution and are largest among top-earning workers in both domestically and foreign-owned firms. However, the sizes of these gaps differ significantly between the two ownership sectors: in domestically owned firms, female workers from the top decile earn around 21% less than male workers; whereas in the foreign-owned firms, this gap amounts to 35%. Interestingly, there are much smaller differences in the shares of women among the top earners in the two ownership sectors: women account for 31% of workers in the top decile of the wage distribution in the domestically owned firms, compared to 28% in the foreign-owned firms.

5 Conclusions

We studied gender pay gaps among workers in domestically and foreign-owned firms in Poland, analysing the differences in the sizes of these gaps. We found that, contrary to most theoretical expectations, gender pay gaps are larger in foreign-owned firms than in domestically owned firms. However, our results also showed that the differences between the two ownership sectors are much smaller than the raw data would suggest.

Moreover, we found that the foreign and the domestic sectors display diverging patterns of gender wage inequality. Women working in the foreign-owned sector earn much less than men based on their personal and workplace characteristics, but the firm-level policies in their workplaces appear to partly reduce this pay disadvantage (as the female employees of foreign-owned companies are paid higher firm-specific wage premia than male employees). In contrast, men working in domestically owned firms receive a higher firm-wage premia than women, which contributes to a female pay disadvantage in the domestic sector. All in all, we found that the average within-firm gender wage gaps are larger in foreign-owned firms than in domestically owned companies, but that the differences

[11] To be precise, around 2% of workers in foreign-owned firms earned the minimum wage, compared to around 20% of workers in domestically owned firms.

between the two ownership sectors are smaller than the economy-wide data would suggest.

In addition to having lower levels of gender wage inequality, the domestic sector is more heterogenous with respect to the sizes of the within-firm gender wage gaps. The right tails of the within-firm gap distribution (representing firms in which women have a large pay disadvantage) are similar for the domestically and foreign-owned companies, but the left tails diverge. In particular, our results showed that almost one-sixth of domestically owned companies pay women more than men (adjusting for individual characteristics), which is more than twice the share of the foreign sector.

This heterogeneity of domestically owned firms also affects the methodology for estimating and interpreting gender wage gaps. Whereas in the foreign sector, the estimates based on male and female characteristics as a reference group yielded similar results, these two approaches yielded more divergent results in the domestic sector. Even though the shares of men and women in total employment in the two ownership sectors are similar, the female and male employees of domestically owned firms tend to be more "dissimilar" in terms of their personal and workplace characteristics (compared to the employees of foreign-owned firms). This is, in our view, the reason why the estimates of gender wage gaps in the domestically owned firms were more susceptible to the choice of the reference group.

More research is needed to identify the factors that could be driving the foreign/domestic firm differences in gender pay gap patterns. First, the question of why there is such a large gender wage gap among the employees of foreign-owned firms remains open. The first step to addressing this question would be to investigate to what extent this gender gap is driven by the sorting channel and gender segregation into lower paying jobs and firms. Second, while there is evidence that firm-level policies can narrow gender wage inequality in other countries, little is known about the ownership divide we identified in our paper. Could it be determined by unions? This is unlikely to be the case in our study. The presence of unions is usually associated with lower within-firm wage inequality, but union density and collective bargaining coverage are low in Poland (and in most other CEE countries) and are even lower in foreign-owned companies than they are in domestically owned firms. Thus, it is unlikely that unions are shaping the equal wage policies of companies in Poland. A portion of the greater heterogeneity of within-firm gender pay gaps in domestically

owned firms could be a legacy of the economic transition, as the domestic companies include both newly created firms and older establishments set up under socialism. The employment and wage policies of these new and old firms are likely to vary more than those of foreign-owned companies. Finally, when we look beyond the mean outcomes, we find that the gender pay gaps are larger in foreign-owned firms than in domestically owned companies at all points of the earnings distribution. However, the gaps are clearly largest among high-paid employees in both ownership sectors.

Our study has its limitations. First, we studied wages earned by men and women, but we were unable to account for non-wage benefits and perks offered by employers. These are likely to differ by sex, and by firm ownership. Second, as our study was limited to paid employees, we were unable to capture the labour incomes of workers with a different employment status (self-employed, those working under special management contracts, etc.). The share of these workers is likely to be highest among the top earners. Third, the use of longitudinal data would allow researchers to better understand the role of firm-level wage premia paid to men and women in domestic and foreign-owned firms.

Appendix

Table 5 Raw and OLS adjusted gender wage gaps in domestically and foreign-owned firms, estimated separately: summary

Ownership	*Raw gender wage gap*	*Models 1a & 1b*	*Models 2a & 2b*	*Models 3a & 3b*	*Models 4a & 4b*
(a) Domestic	10.0	15.4	10.9	10.6	10.5
(b) Foreign	27.3	23.9	17.9	17.9	17.9
Model specification					
Individual characteristics	–	Yes	Yes	Yes	Yes
Job characteristics	–	Yes	Yes	Yes	Yes
Firm characteristics	–	Yes	Yes	No	No
Co-workers characteristics	–	No	Yes	No	No
Firm fixed effects	–	No	No	Yes	Yes
Common support	No	No	No	No	Yes

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: The full set of estimates is available in Table 7 (models 1a, 2a, 3a, 4a) and Table 8 (models 1b, 2b, 3b, 4b)

Table 6 Regression results: gender wage gap in domestic and foreign-owned firms

	Logarithm of wage (OLS, Model 1)	*Logarithm of wage (OLS, Model 2)*	*Logarithm of wage (OLS, Model 3)*	*Logarithm of wage (OLS, Model 4)*
Female	-0.164***	-0.107***	-0.116***	-0.117***
	(0.005)	(0.005)	(0.002)	(0.002)
Foreign	0.293***	0.232***	0.668***	0.666***
	(0.013)	(0.011)	(0.052)	-0.053***
Female x foreign	-0.085***	-0.087***	-0.056***	(0.003)
	(0.015)	(0.012)	(0.003)	0.027***
Age	0.030***	0.027***	0.027***	(0.000)
	(0.001)	(0.001)	(0.000)	-0.000***
Age squared	-0.000***	-0.000***	-0.000***	(0.000)
	(0.000)	(0.000)	(0.000)	0.010***
Education: basic vocational (base: primary)	0.015**	-0.000	0.011***	(0.002)
	(0.006)	(0.010)	(0.002)	0.045***
Education: secondary (base: primary)	0.083***	0.048***	0.045***	-0.053***
	(0.007)	(0.011)	(0.002)	(0.002)
Education: tertiary (base: primary)	0.334***	0.193***	0.199***	0.196***
	(0.011)	(0.010)	(0.003)	(0.003)
Experience	0.003***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Part-time dummy	0.010	0.004	0.003	0.002
	(0.007)	(0.007)	(0.002)	(0.002)
Fixed-term contract dummy	-0.118***	-0.094***	-0.085***	-0.083***
	(0.007)	(0.006)	(0.002)	(0.002)
Tenure	0.658***	0.655***	0.005***	0.005***
	(0.035)	(0.034)	(0.000)	(0.000)
Firm size: more than 250 workers (base: 51–250)	0.374***	0.310***		
	(0.032)	(0.034)		
Firm size: 10–50 workers (base: 51–250)	0.253***	0.212***		
	(0.031)	(0.033)		
Collective bargaining	0.088***	0.065*		
	(0.031)	(0.033)		
Share of women		-0.002***		
		(0.000)		
Share of workers with tertiary education		0.006***		
		(0.000)		
Share of workers under age 35		-0.002***		
		(0.000)		

(*continued*)

Table 6 (continued)

	Logarithm of wage (OLS, Model 1)	*Logarithm of wage (OLS, Model 2)*	*Logarithm of wage (OLS, Model 3)*	*Logarithm of wage (OLS, Model 4)*
Share of workers aged 55 or older		-0.002***		
		(0.000)		
Other controls:				
occupation dummies	Yes	Yes	Yes	Yes
NACE dummies	Yes	Yes	No	No
Firm fixed effects	No	No	Yes	Yes
Observations	343,143	343,143	343,143	333,661 (common support)
R-squared/ pseudo-R-squared	0.537	0.576	0.757	0.760

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: Models with an intercept. Standard errors in parentheses. In models 1 and 2, standard errors clustered at the firm level. In models 3 and 4, cluster-robust variance estimator used. ***p<0.01, **p<0.05, *p<0.1

Table 7 Regression results: gender wage gap in domestically owned firms

	Logarithm of wage (OLS, Model 1a)	*Logarithm of wage (OLS, Model 2a)*	*Logarithm of wage (OLS, Model 3a)*	*Logarithm of wage (OLS, Model 4a)*
Female	-0.154***	-0.109***	-0.106***	-0.105***
	(0.005)	(0.003)	(0.002)	(0.002)
Age	0.018***	0.017***	0.018***	0.018***
	(0.001)	(0.001)	(0.000)	(0.000)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Education: basic vocational (base: primary)	0.019***	0.015***	0.021***	0.021***
	(0.006)	(0.006)	(0.002)	(0.002)
Education: secondary (base: primary)	0.084***	0.065***	0.055***	0.055***
	(0.006)	(0.006)	(0.002)	(0.002)
Education: tertiary (base: primary)	0.283***	0.179***	0.182***	0.176***
	(0.008)	(0.007)	(0.003)	(0.003)
Experience	0.003***	0.003***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Part-time dummy	0.019**	0.013*	0.015***	0.013***
	(0.007)	(0.007)	(0.002)	(0.002)

(*continued*)

Table 7 (continued)

	Logarithm of wage (OLS, Model 1a)	*Logarithm of wage (OLS, Model 2a)*	*Logarithm of wage (OLS, Model 3a)*	*Logarithm of wage (OLS, Model 4a)*
Fixed-term contract dummy	-0.107*** (0.006)	-0.088*** (0.006)	-0.075*** (0.002)	-0.075*** (0.002)
Tenure	0.005*** (0.000)	0.005*** (0.000)	0.005*** (0.000)	0.005*** (0.000)
Firm size: more than 250 workers (base: 51–250)	0.096*** (0.015)	0.092*** (0.014)		
Firm size: 10–50 workers (base: 51–250)	-0.139*** (0.009)	-0.131*** (0.009)		
Collective bargaining	0.033*** (0.009)	0.037*** (0.009)		
Share of women		-0.001*** (0.000)		
Share of workers with tertiary education		0.006*** (0.000)		
Share of workers under age 35		-0.001*** (0.000)		
Share of workers aged 55 or older		-0.000 (0.000)		
Other controls:				
Occupation dummies	Yes	Yes	Yes	Yes
NACE dummies	Yes	Yes	No	No
Firm fixed effects	No	No	Yes	Yes
Observations	222,203	222,203	222,203	215,944 (common support)
R-squared/ pseudo-R-squared	0.447	0.480	0.730	0.732

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: Models with an intercept. Standard errors in parentheses. In models 1a and 2a, standard errors clustered at the firm level. In models 3a and 4a, cluster-robust variance estimator used. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Table 8 Regression results: gender wage gap in foreign-owned firms

	Logarithm of wage (OLS, Model 1b)	*Logarithm of wage (OLS, Model 2b)*	*Logarithm of wage (OLS, Model 3b)*	*Logarithm of wage (OLS, Model 4b)*
Female	-0.239***	-0.179***	-0.179***	-0.179***
	(0.012)	(0.008)	(0.003)	(0.003)
Age	0.049***	0.043***	0.043***	0.043***
	(0.003)	(0.003)	(0.001)	(0.001)
Age squared	-0.001***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Education: basic vocational (base: primary)	-0.040	-0.066*	-0.047***	-0.049***
	(0.027)	(0.034)	(0.004)	(0.004)
Education: secondary (base: primary)	0.059**	0.006	0.007*	0.006*
	(0.030)	(0.038)	(0.004)	(0.004)
Education: tertiary (base: primary)	0.380***	0.209***	0.204***	0.205***
	(0.023)	(0.030)	(0.005)	(0.005)
Experience	0.002*	0.002**	0.002***	0.002***
	(0.001)	(0.001)	(0.000)	(0.000)
Part-time dummy	0.013	0.006	-0.010**	-0.009**
	(0.017)	(0.018)	(0.004)	(0.004)
Fixed-term contract dummy	-0.158***	-0.120***	-0.103***	-0.101***
	(0.017)	(0.016)	(0.004)	(0.004)
Tenure	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.000)	(0.000)
Firm size: more than 250 workers (base: 51–250)	-0.036*	0.025		
	(0.020)	(0.016)		
Firm size: 10–50 workers (base: 51–250)	0.023	-0.010		
	(0.025)	(0.021)		
Collective bargaining	0.010	0.018		
	(0.022)	(0.018)		
Share of women		-0.003***		
		(0.000)		
Share of workers with tertiary education		0.007***		
		(0.000)		
Share of workers under age 35		-0.003***		
		(0.001)		
Share of workers aged 55 or older		-0.008***		
		(0.001)		
Other controls:				

(*continued*)

Table 8 (continued)

	Logarithm of wage (OLS, Model 1b)	*Logarithm of wage (OLS, Model 2b)*	*Logarithm of wage (OLS, Model 3b)*	*Logarithm of wage (OLS, Model 4b)*
Occupation dummies	Yes	Yes	Yes	Yes
NACE dummies	Yes	Yes	No	No
Firm fixed effects	No	No	Yes	Yes
Observations	120,940	120,940	120,940	117,717 (common support)
R-squared/ pseudo-R-squared	0.591	0.637	0.748	0.748

Source: Own calculations based on the Structure of Wages and Salaries by Occupations 2014 data

Notes: Models with an intercept. Standard errors in parentheses. In models 1b and 2b, standard errors clustered at the firm level. In models 3b and 4b, cluster-robust variance estimator used. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

References

Arrow, K. (1973). The Theory of Discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets* (pp. 3–33). Princeton University Press.

Bayard, K., Hellerstein, J., Neumark, D., & Troske, K. (2003). New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data. *Journal of Labor Economics, 21*(4), 887–922.

Becker, G. S. (1957 [1971]). *The Economics of Discrimination* (2nd ed.). University of Chicago Press.

Black, S. E., & Brainerd, E. (2004). Importing Equality? The Impact of Globalization on Gender Discrimination. *Industrial & Labor Relations Review, 57*(4), 540–559.

Black, S. E., & Strahan, P. E. (2001). The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry. *American Economic Review, 91*(4), 814–831.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources, 8*, 436–455.

Bøler, E. A., Javorcik, B., & Ulltveit-Moe, K. H. (2018). Working Across Time Zones: Exporters and the Gender Wage Gap. *Journal of International Economics, 111*, 122–133.

Braunstein, E., & Brenner, M. (2007). Foreign Direct Investment and Gendered Wages in Urban China. *Feminist Economics, 13*(3–4), 213–237.

Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *Quarterly Journal of Economics, 131*(2), 633–686.

Central Statistical Office (CSO). (2015). *Economic activity of entities with foreign capitalin 2014*. Warsaw, Poland.

Christofides, L. N., Polycarpou, A., & Vrachimis, K. (2013). Gender Wage Gaps, 'Sticky Floors' and 'Glass Ceilings' in Europe. *Labour Economics, 21*, 86–102.

Cotton, J. (1988). On the Decomposition of Wage Differentials. *The Review of Economics and Statistics, 70*, 236–243.

Duncan, O. D., & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review, 20*(2), 210.

Earle, J. S., Telegdy, Á., & Antal, G. (2018). Foreign Ownership and Wages: Evidence from Hungary, 1986–2008. *ILR Review, 71*(2), 458–491.

Felfe, C. (2012). The Motherhood Wage Gap: What About Job Amenities? *Labour Economics, 19*(1), 59–67.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica, 77*(3), 953–973.

Fortin, N. (2008). The Gender Wage Gap Among Young Adults in The United States. *The Journal of Human Resources, 43*(4), 884–918. University of Wisconsin Press.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review, 104*(4), 1091–1119.

Goraus, K., Tyrowicz, J., & Van der Velde, L. (2017). Which Gender Wage Gap Estimates To Trust? A Comparative Analysis. *Review of Income and Wealth, 63*(1), 118–146.

Heinze, A., & Wolf, E. (2010). The Intra-Firm Gender Wage Gap: A New View on Wage Differentials Based on Linked Employer–Employee Data. *Journal of Population Economics, 23*(3), 851–879.

Hijzen, A., Martins, P. S., Schank, T., & Upward, R. (2013). Foreign-Owned Firms around the World: A Comparative Analysis of Wages and Employment at the Micro-Level. *European Economic Review, 60*, 170–188.

Javdani, M. (2015). Glass Ceilings or Glass Doors? The Role of Firms in Male-Female Wage Disparities. *Canadian Journal of Economics/ Revue canadienne d'économie, 48*(2), 529–560.

Killewald, A., & Bearak, J. (2014). Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression. *American Sociological Review, 79*(2), 350–357.

Kodama, N., Javorcik, B. S., & Abe, Y. (2018). Transplanting Corporate Culture across International Borders: Foreign Direct Investment and Female Employment in Japan. *The World Economy, 41*(5), 1148–1165.

Kunze, A. (2008). Gender Wage Gap Studies: Consistency and Decomposition. *Empirical Economics, 35*, 63–76.

Lechner, M., & Strittmatter, A. (2019). Practical Procedures to Deal with Common Support Problems in Matching Estimation. *Econometric Reviews, 38*(2), 193–207.

Li, L., & Dong, X.-Y. (2011). Economic Transition and the Gender Earnings Gap in Chinese Industry: The Role of Firm Characteristics. *Contemporary Economic Policy, 29*(1), 67–87.

Liu, P.-W., Meng, X., & Zhang, J. (2000). Sectoral Gender Wage Differentials and Discrimination in the Transitional Chinese Economy. *Journal of Population Economics, 13*(2), 331–352.

Lovasz, A. (2008). *Competition and the Gender Wage Gap: New Evidence from Linked Employer-Employee Data in Hungary 1986–2003.* Budapest Working Papers on the Labour Market 0804. Institute of Economics, Centre for Economic and Regional Studies.

Magda, I., & Cukrowska-Torzewska, E. (2019). Do Women Managers Lower Gender Pay Gaps? Evidence from Public and Private Firms. *Feminist Economics, 25*(4), 185–210.

Matysiak, A., Słoczyński, T., & Baranowska, A. (2010). Kobiety i Mężczyźni na Rynku Pracy,[w:] Zatrudnienie w Polsce 2008. Praca w cyklu życia, red. M. Bukowski, Warszawa.

Maurer-Fazio, M., & Hughes, J. (2002). The Effects of Market Liberalization on the Relative Earnings of Chinese Women. *Journal of Comparative Economics, 30*(4), 709–731.

Meng, X. (2004). Gender Earnings Gap: the Role of Firm Specific Effects. *Labour Economics, 11*(5), 555–573.

Meng, X., & Meurs, D. (2004). The Gender Earnings Gap: Effects of Institutions and Firms—A Comparative Study of French and Australian Private Firms. *Oxford Economic Papers, 56*(2), 189–208.

Nekby, L. (2003). Gender Differences in Rent Sharing and its Implications for the Gender Wage Gap, Evidence from Sweden. *Economics Letters, 81*(3), 403–410.

Neumark, D. (1988). Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *The Journal of Human Resources, 23*(3), 279–295.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*(3), 693–709.

Oaxaca, R. L., & Ransom, M. R. (1994). On Discrimination and the Decomposition of Wage Differentials. *Journal of Econometrics, 61*(1), 5–21.

Oostendorp, R. H. (2009). Globalization and the Gender Wage Gap. *The World Bank Economic Review, 23*(1), 141–161.

Reilly, K. T., & Wirjanto, T. S. (1999). Does More Mean Less? The Male/Female Wage Gap and the Proportion of Females at the Establishment Level. *Canadian Journal of Economics, 32*(4), 906–929.

Reimers, C. W. (1983). Labor Market Discrimination Against Hispanic and Black Men. *The Review of Economics and Statistics, 65*, 570–579.

Rickne, J. (2012). Gender and Work Compensation in China's Industrial Sector. *Review of Income and Wealth, 58*(2), 307–329.

Seguino, S., & Grown, C. (2006). Gender Equity and Globalization: Macroeconomic Policy for Developing Countries. *Journal of International Development, 18*(8), 1081–1104.

Słoczyński, T. (2020). Average gaps and Oaxaca–Blinder decompositions: A cautionary tale about regression estimates of racial differences in labor market outcomes. *ILR Review, 73*(3), 705–729.

Song, J., Price, D. J., Guyenen, F., Bloom, N., & von Wachter, T. (2019). Firming up Inequality. *The Quarterly Journal of Economics, 134*(1), 1–50. Published 25 October 2018.

Statistics Poland. (2017). *Economic Activity of Entities with Foreign Capital in 2016*. Warsaw, Poland.

Strawiński, P., Majchrowska, A., & Broniatowska, P. (2018). Occupational Segregation and Wage Differences: the Case of Poland. *International Journal of Manpower, 39*(3), 378–397.

Tyrowicz, J., & Smyk, M. (2019). Wage Inequality and Structural Change. *Social Indicators Research, 141*(2), 503–538.

Vahter, P., & Masso, J. (2019). The Contribution of Multinationals to Wage Inequality: Foreign Ownership and the Gender Pay Gap. *Review of World Economics, 155*(1), 105–148.

Zulfiu-Alili, M. (2014). Inward Foreign Direct Investment and Wage Inequality in Macedonia. *Eastern European Economics, 52*(5), 56–86.

Zweimueller, M., Winter-Ebmer, R., & Weichselbaumer, D. (2008). Market Orientation and Gender Wage Gaps: an International Study. *KYKLOS, 61*(4), 615–635.

An International Organization's Approach to Labour Migration for Human Development

Laura Bartolini

Labour migration[1] can be broadly defined as the movement of persons from their origin location of residence to another location within the same country (internal) or in another country (international) for the purpose of

[1] The author acknowledges the multifaceted dimensions of migration and the existence of distinct migration pathways—student mobility, family reunification, entrepreneur or investor schemes, resettlement, relocation and complementary pathways for those fleeing persecution war and conflict, to name some of them. This chapter focuses on labour migration only.

The opinions expressed in this chapter are those of the author and do not necessarily reflect the views of the International Organization for Migration (IOM). This article was closed in September 2021.

L. Bartolini (✉)
International Organization for Migration, Regional Office for South-Eastern Europe, Eastern Europe and Central Asia, Vienna, Austria
e-mail: lbartolini@iom.int

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_9

employment. Migration intentions can be seen as the result of balancing the potential gains and costs of moving: if the potential gains exceed the costs of relocating, mobility can be preferred to immobility. Of course, aspirations and capabilities to move (Carling, 2002; Schewel, 2019; de Haas, 2021) do not always coincide. And of course, what one should include in the calculation of "gains" and "costs," at the personal household and societal level, is debatable and debated. Yet, that "migration is generally beneficial and that greater openness would create large global economic gains" (Dustmann & Preston, 2019) is now more and more accepted among economists in academia,[2] but faces increasing hesitations or resistance in most countries—and especially in the main immigration countries in the world—which on the contrary put in place efforts to manage, hide, limit or stop immigration to their societies altogether. Cleary, considerations about migrations are and should not be grounded on economic terms alone, and the recognition of the distributive economic effects as well as of the many non-economic factors at play is needed for the global governance(s) of migrations to be grounded on the respect of the rights, dignity and well-being of migrants, and to be more satisfactory and efficient in economic terms for all those involved.

In this regard, the International Organization for Migration (IOM) is committed since its establishment in 1951 to the principle that humane and orderly migration benefits migrants and society. As an intergovernmental organization, IOM has offered support to governments over the decades in meeting the challenges of managing all forms of migration in accordance with international law, while supporting migrants across the world, particularly those in situation of vulnerability and in emergency situations. From its roots as an operational logistics agency, IOM has broadened its scope of action and restructured itself in line with the shifting dynamics of migrations over the decades, recently joining the United Nations system in 2016 and becoming the Coordinator and Secretariat for the UN Migration Network established after the UN adoption of the Global Compact for Safe, Orderly and Regular Migration (GCM) (UN General Assembly, 2018) in late 2018.

IOM's most recent internal developments in terms of understanding of and approaches to labour mobility and human development—in line with

[2] Letting alone the human rights' considerations that are more frequently discussed by scholars of other disciplines.

the existing international legal framework on the matter[3]—can be retrieved from the actual implementation of some key global initiatives and programmes, from its most recent institutional documents (Migration Governance Framework endorsed by IOM Member States in 2015 and the five-year Strategic Vision of 2019) and from those stemming from its participation in the drafting process of the GCM.

Here, labour migration is conceptualized as potentially beneficial for origin countries as it can help reducing unemployment and underemployment, and it can bring important resources from abroad to low- and middle-income countries in terms of remittances, skills' transfer, diasporas' engagement and the creation of new commercial and business activities. In main destination countries, labour migration can help addressing specific labour shortages. At the same time, migration can bring improvements in terms of income, skills, competencies, living conditions for migrants, their families and their communities. Yet, a number of governance challenges are acknowledged for all actors and level involved. For origin countries, labour emigration can bring potential skills shortages (the so-called brain drain) and changes to social structures in families and communities left behind. At destination, the availability of migrant workforces can have important distributive impacts in terms of employment and remuneration of local workers in specific sectors and occupations, potentially bringing more disarticulation of local labour markets and social tensions especially among workers in low-skilled occupations. What is more, from an operational and humanitarian perspective, there are challenges associated with the protection of migrant workers' rights at each step of the migration process, at origin, transit and destination. Over the years, IOM has developed programmes and capacity-building activities for local and national stakeholders to identify and be able to contrast

[3] The international normative frameworks protecting the rights of migrant workers can be found in numerous international human rights treaties that apply to all persons and in treaties specifically focused on migrants as the International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families of 1990 (entered into force in 2003 and still to be signed many UN member states). Criminal aspects related to migration are addressed in transnational criminal law (e.g. see the UN Convention against Transnational Organized Crime, the Protocols against Smuggling of Migrants by Land, Sea and Air, and to Prevent, Suppress and Punish Trafficking in Persons, especially Women and Children of 2000), while international labour standards for all persons and for migrants specifically are contained in the core International Labour Organization (ILO) conventions. In 2016, ILO has also created the Multilateral Framework on Labour Migration (ILO, 2006) that contains non-binding principles and guidelines for a rights-based approach to labour migration.

deceptive and non-transparent recruitment procedures, debt bondage, labour exploitation and to protect migrants in vulnerable situations—especially those who migrate irregularly—from smuggling and human trafficking networks. Once at destination, migrant workers can experience discrimination, diminished access to rights and services compared to local workers, the risk of detention and deportation connected with their employment and residence status, while it is generally more difficult for them to access existing judicial remedies. Moreover, IOM acknowledges the specific conditions of migrant women and children in the labour markets given by the intertwined effect of their gender, age, immigration and employment status (IOM, 2018).

In line with this overall understanding, IOM labour mobility and human development agenda unfolds under several strands of work "offering policy and technical advice to national governments, supporting the development of policies legislation and administrative structures that promote efficient, effective and transparent labour migration flows, assisting governments to promote safe labour migration practices for their nationals, facilitating the recruitment of workers, including pre-departure training and embarkation preparedness, promoting the integration of labour migrants in their new workplace and society"[4] and in the implementation of these activities IOM engages with migrants, their families and communities, with local, national and regional bodies and governments, and more recently also with representatives of the private sector (employers' associations and unions). Among the more notable examples of this comprehensive approach, one can mention the global International Recruitment Integrity System (IRIS)[5]—IOM's flagship initiative to promote ethical recruitment which is also referred under Objective 6 of the GCM—as well as other initiatives aimed at promoting intra-regional highly skilled migration (e.g. in the case of ASEAN), supporting free movement of persons within existing regional organizations (ECOWAS) as well as the establishment of bilateral labour migration agreements among countries with already well-established migration corridors, working with governments in establishing skills mobility partnerships in line with the most updated evidence as collected by the Center for Global Development[6] (Adhikari et al., 2021).

[4] See: https://www.iom.int/labour-migration
[5] See: https://iris.iom.int/
[6] See: https://gsp.cgdev.org/

While the long-term effects of the COVID-19 pandemic on the structure of global and local labour markets are still to be seen in full, practically all UN agencies and international organizations agree (UN, 2020; Bendon et al., 2021; ILO, 2021a, 2021b; UNDP, 2021; FAO, 2021) in claiming that migrant workers have been disproportionately impacted by the COVID-19-related crises, given the specific occupations, sectors and locations where they are more represented, with important cascade effects on their families and communities in their residence and origin countries.

And yet, unlike what some have argued during the first phase of the pandemic, 2020 has not been the start of a new era of zero migration. While it is true that some of the more traditional dynamics, norms, policies and practices of migrations globally have been disrupted since the outbreak of COVID-19, still international migration remains as an important feature of global labour markets and societies.

Among the many challenges posed or exposed by the pandemic under various domains of our societies, one surely is the establishment of more coherent, predictable and equitable migration and border management systems to mitigate the unequal effects of protracted and differentiated border closures and mobility restrictions, to avoid that these turn into additional implicit mechanisms by which main destination countries skim and control immigration and that more individuals are forced to resort to irregular migration channels. As recently argued (Freier, 2020), pathways to regular migration and regularization mechanisms, as those also advocated by IOM, are needed not only on economic and rights-based grounds, to protect individuals along the migration process, but also from a public and global health perspective. With only 3.6% of the world population living outside its country of origin in 2020 (United Nations, 2019), the desire to stay is still far more common globally than one would expect given the persisting and increasing disparities in wealth, well-being and security around the world. For those who aspire to move, policies that make them able to do so regularly will reduce risks and costs for them and for the countries involved and enhance human security overall.

References

Adhikari, S., Clemens, M., & Dempster, H. (2021). *Why Europe Should Build Legal Migration Pathways with Nigeria*. Center for Global Development. Retrieved September 7, 2021, from https://www.cgdev.org/publication/why-europe-should-build-legal-migration-pathways-nigeria

Bendon, M., Batalova, J., Davidoff-Gore, S., & Schmidt, T. (2021). *COVID-19 and the State of Global Mobility in 2020* (p. 63). Migration Policy Institute and International Organization for Migration. Retrieved September 7, 2021, from https://publications.iom.int/books/covid-19-and-state-global-mobility-2020

Carling, J. (2002). Migration in the Age of Involuntary Immobility: Theoretical Reflections and Cape Verdean Experiences. *Journal of Ethnic and Migration Studies, 28*(1), 5–42. https://doi.org/10.1080/13691830120103912

de Haas, H. (2021). A Theory of Migration: The Aspirations-Capabilities Framework. *Comparative Migration Studies, 9*(1), 8. https://doi.org/10.1186/s40878-020-00210-4

Dustmann, C., & Preston, I. P. (2019). Free Movement, Open Borders, and the Global Gains from Labor Mobility. *Annual Review of Economics, 11*(1), 783–808. https://doi.org/10.1146/annurev-economics-080218-025843

FAO. (2021). *Seasonal Migration in Europe and Central Asia in the Context of the COVID-19 Pandemic*. FAO. Retrieved July 9, 2021, from http://www.fao.org/documents/card/en/c/cb5718en/

Freier, L. F. (2020). *COVID-19 and Rethinking the Need for Legal Pathways to Mobility: Taking Human Security Seriously* (p. 11). International Organization for Migration. Retrieved August 9, 2021, from https://publications.iom.int/system/files/pdf/rethinking-the-need-for-legal.pdf

ILO. (2006). *The Multilateral Framework on Labour Migration*. International Labour Organization. Retrieved September 7, 2021, from http://www.ilo.org/global/topics/labour-migration/publications/WCMS_178672/lang%2D%2Den/index.htm

ILO. (2021a). *World Employment and Social Outlook: Trends 2021* (p. 164). International Labour Office. https://www.ilo.org/wcmsp5/groups/public/%2D%2D-dgreports/%2D%2D-dcomm/%2D%2D-publ/documents/publication/wcms_795453.pdf

ILO. (2021b). *World Social Protection Report 2020–22: Social Protection at the Crossroads—In Pursuit of a Better Future*. International Labour Office. https://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/@publ/documents/publication/wcms_817572.pdf

IOM. (2018). *Expanding Labour Mobility Channels* (p. 10). IOM. https://www.iom.int/sites/g/files/tmzbdl486/files/our_work/ODG/GCM/IOMThematicPaper-ExpandingLabourMobilityChannels.pdf

Schewel, K. (2019). Understanding Immobility: Moving Beyond the Mobility Bias in Migration Studies. *International Migration Review, 54*(2), 328–355. https://doi.org/10.1177/0197918319831952

UN. (2020). *Policy Brief: COVID-19 and People on the Move* (p. 25).United Nations. Retrieved July 9, 2021, from https://unsdg.un.org/sites/default/files/2020-06/SG-Policy-Brief-on-People-on-the-Move.pdf

UN General Assembly. (2018). *Global Compact for Safe, Orderly and Regular Migration—Resolution A/RES/73/195*. United Nations. https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/73/195

UNDP. (2021). *Socio-Economic Impact of COVID-19 on Women Migrant Workers*. United Nations Development Programme. Retrieved September 7, 2021, from https://www.undp.org/publications/socio-economic-impact-covid-19-women-migrant-workers

United Nations. (2019). *World Population Prospects 2019—Online Edition. Rev. 1*. UN Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp/Download/Standard/Population/

Deglobalization and Labour: A New Era?

Peter A. G. van Bergeijk and Rolph van der Hoeven

Since the mid-1990s globalization has increasingly been associated with reduced demand for low-skilled jobs in industrialized countries and with an increasing tension for fair labour standards such as minimum wages and workplace regulations in developing countries, but also in industrialized countries. However, in a flat world, so it was argued by the proponents of globalization, regulatory competition to improve labour standards would lead to lower growth (Friedman, 2005). Distance was dead and therefore non- or low-skilled jobs in the advanced Western market economies would be lost to other countries, but the presumed economic growth in advanced countries would be able to compensate the losers from globalization. The policy prescription was clear, but the results were not. Workers in industrialized countries and social activist in developing countries actively pleaded for deglobalization policies. That wish may have become true: after the Great Recession in 2008 international trade declined and protectionism increased. But has it brought a better world for labour in general and

P. A. G. van Bergeijk (✉) • R. van der Hoeven
The International Institute of Social Studies (ISS) of Erasmus University Rotterdam, The Hague, Netherlands
e-mail: bergeijk@iss.nl; hoeven@iss.nl

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_10

more specifically after the outbreak of the COVID-19 pandemic? How will all this play out for labour in a post-pandemic world.?

It is important to observe that the second wave of globalization already had come to an end at the time Walden Bello produced his influential *Deglobalization: Ideas for a new world economy* in Bello (2005). The culmination point of globalization seems to have been around 2007/8 just before the start of the financial crisis. A phase of deglobalization was set in motion driven by a global trade slow down, increasing trade uncertainty and geopolitically inspired trade destruction by the former US president and the current UK prime minister (van Bergeijk, 2019). The SARS-CoV-2 virus therefore did hit a global system with underlying conditions of lacklustre global cooperation (van Bergeijk, 2021). It also exposed a longer trend that drove the income share of the super-rich up and the global poor down: The labour share shrunk as the capital labour ratio increased, asset price inflation fed housing prices and stock markets alike and white-collar jobs could by and large be maintained working over the internet from home. Between countries, income disparities increased. Within countries, inequality increased everywhere, in low-, middle- and high-income countries.

According to the ILO (2021) poor workers are becoming poorer as some 600 million people work in sectors which are hardest hit and that pay poorly. The marginalized poor and informal sector workers live and work under conditions that are conducive to contamination, the work places are over-crowded, housing is poor often without access to hygienic facilities and lockdowns are not a realistic tool since their livelihoods are threatened. On top of that, the generation gap is increasing, with a greater number of younger workers being excluded from the labour market and having to work under precarious conditions, while relatively privileged workers are better sheltered from the COVID-19 economic outfall (Fig. 1).

The fall out of COVID-19 reinforces the ongoing trend of deglobalization. Should we therefore embrace globalization again as happened at the end of the twentieth century? That would be an unwise and too fast conclusion. As Stiglitz (2002) did argue two decades ago, the critics of globalization were concerned by unfettered globalization that was not managed, neither at international nor at national level. The observed increasing inequality is a consequence of the fact that global shocks (both positive and negative) and the absence of global governance in the end always favour the (internationally) mobile factors of production, typically capital

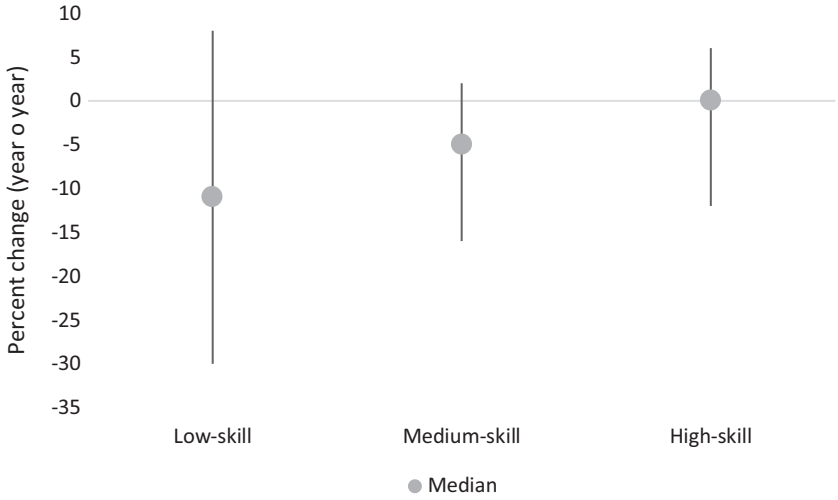

Fig. 1 Change in employment by level of skills (2020Q2). (Source: calculations based on data underlying ILO (2021, Figure B2))

(owners) and the better educated (Stiglitz, 2002, 2018. van der Hoeven, 2017).

So while globalization and deglobalization are different phenomena that look like symmetric concepts, the impact of the shocks on labour is similar. It is therefore important to consider under which national and international rules globalization and deglobalization can be managed, along similar lines as Nayyar (2020) has argued in *Governing Globalization, Issues and Institutions*. Our point is that deglobalization needs the same (global governance) for the same reason (the impact on livelihoods). The World Commission on the Social Dimensions of Globalization (2004) succinctly stated:

> *We believe the dominant perspective on globalization must shift more from a narrow preoccupation with markets to a broader preoccupation with people. Globalization must be brought from the high pedestal of corporate board rooms and cabinet meetings to meet the needs of people in the communities in which they live. The social dimension of globalization is about jobs, health and education—but it goes far beyond these. It is the dimension of globalization which people experience in their daily life and work: the totality of their aspirations for democratic participation and material prosperity. A better globalization is the*

> *key to a better and secure life for people everywhere in the twenty-first century. We also propose a process by which such a perspective can be realized at all levels, beginning with empowered local communities and improved and more accountable national governance; fair global rules applied fairly; and global institutions that are more pro-people.* (p. vii)

Van der Hoeven and Vos (2021) argue that the global consequences of COVID-19 could have been less severe, and that much of the increase in global poverty could have been prevented, if the international financial system had been fit for purpose and the supply of vaccines would have been global and not captured by the developed countries.

The current context of a post-pandemic recovery provides a unique opportunity for a new approach that needs to be different from the dominating financial globalization and unchecked opening up of economies without much consideration of potential drawbacks that characterized the second wave of globalization at the end of the last century. This new phase should be characterized by recognizing the political and societal harm of a cleavage in the labour market and of growing inequality between capital and labour, without retreating into protectionism as populist leaders in the North and the South proclaim.

Gallagher and Kozul-Wright (2020) call for a 'New Multilateralism for Shared Prosperity', that recognizes the fault lines of the second wave of globalization that Rodrik (2019) has labelled 'the "hyper-globalism," under which the priorities of the global economy receive precedence over the priorities of the home economy. According to this model for the international system, countries must maximally open their economies to foreign trade and investment, regardless of the consequences for their growth strategies or social models'. 'Ultimately, a healthy and sustainable world trade regime would be one of "peaceful economic coexistence," in which different economic systems prosper side by side rather than being pressured to conform to a single mold favored by international corporations' according to Rodrik. Deglobalization is not the answer to hyperglobalism, of course.

Rodrik and Stantcheva (2020) call for a new economic order requiring an explicit quid pro quo between private firms and public authorities. 'To prosper, firms need a reliable and skilled workforce, good infrastructure, an ecosystem of suppliers and collaborators, easy access to technology, and a sound regime of contracts and property rights. Most of these are provided through public and collective action, which is the government's side

of the bargain'. Governments, of course, must ensure that firms are incentivised to internalize the external effects of their decisions to produce for their communities and societies. 'And firms must live up to their side of the bargain—not as a matter of corporate social responsibility, but as part of an explicit regulatory and governance framework'. Importantly this new approach should no longer separate pro-growth policies and social policies. Faster economic growth requires disseminating new technologies and productive opportunities among smaller firms and wider segments of the labour force, rather than confining the use of new technologies and neue kombinationen and their benefits to the elite the target should be to decrease inequalities and insecurities. Moreover, Rodrick and Stantcheva (2020) are spot on when they argue that growth and social agendas are one and the same, but their focus on national agendas misses the point that the pressing problems are global and require a global answer. We need to move from macro-economics to earth economics. The post-pandemic world despite the current phase of deglobalization remains connected as never before. The pandemic has forcefully made that point and suggests urgent action to reduce the inequalities that are the breeding ground of the next pandemic.

Key in the post-pandemic policies would be a recognition of the plight of labour especially the growing group of precarious workers in both developed and developing countries, because it is they who have lost out most.

References

Bello, W. (2005). *Deglobalization: Ideas for a New World Economy*. Zed Books.

Friedman, T. L. (2005). *The World is Flat*. Penguin.

Gallagher, K. P., & Kozul-Wright, R. (2020). *A New Multilateralism for Shared Prosperity*. Boston University and UNCTAD.

ILO. (2021). *ILO Monitor: COVID-19 and the World of Work* (7 ed.).

Nayyar, D. (2020). *Governing Globalization, Issues and Institutions*. Oxford University Press.

Rodrik, D. (2019). *The Need for a Global Trade Makeover*. Project Syndicate.

Rodrik, D., & Stantcheva, S. (2020). *The Post-Pandemic Social Contract*. Project Syndicate.

Stiglitz, J. E. (2002). *Globalization and Its Discontents*. Norton.

Stiglitz, J. E. (2018). *Globalization and its Discontents Revisited. Anti-globalization in the Era of Trump*. Norton.

van Bergeijk, P. A. G. (2019). *Deglobalization 2.0. Trade and Openness During the Great Depression and the Great Recession*. Cheltenham.

van Bergeijk, P. A. G. (2021). *Pandemic Economics*. Edward Elgar.

van der Hoeven, R. (2017). *Can the SDGs stem rising income inequality in the world?* Chap. 10 in van Bergeijk and van der Hoeven (Eds.) (2017), *Sustainable Development Goals and Income Inequality*. Edward Elgar,

van der Hoeven, R, & Vos, R. (2021). Reforming the International Financial and Fiscal System for better COVID-19 and Post-Pandemic Crisis Responsiveness. Chap. 2 in E. Papyrakis (Ed.), *COVID19 and International Development*. Springer/Palgrave

World Commission on the Social Dimension of Globalization. (2004). *A fair Globalization, Creating Opportunities for All*. ILO.

PART II

Jobs and Technological Change

Technological Change and the Future of Work

Raul Ramos, Gianluca Ferrittu, and Pedro Goulart

1 Nearly 30 Years of Living (and Working) with the Internet

The internet has changed our lives. It has represented a revolution in the way we communicate. In fact, information technologies have brought fundamental change throughout society, driving it forward from the

R. Ramos (✉)
Department of Econometrics, Statistics and Applied Economics, AQR-IREA, University of Barcelona, Barcelona, Spain
e-mail: rramos@ub.edu

G. Ferrittu
Lisbon School of Economics and Management - ISEG, Universidade de Lisboa, Lisbon, Portugal
e-mail: gferrittu@phd.iseg.ulisboa.pt

P. Goulart
CAPP, Institute of Social and Political Sciences, Universidade de Lisboa, Lisbon, Portugal
e-mail: pgoulart@edu.ulisboa.pt

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_11

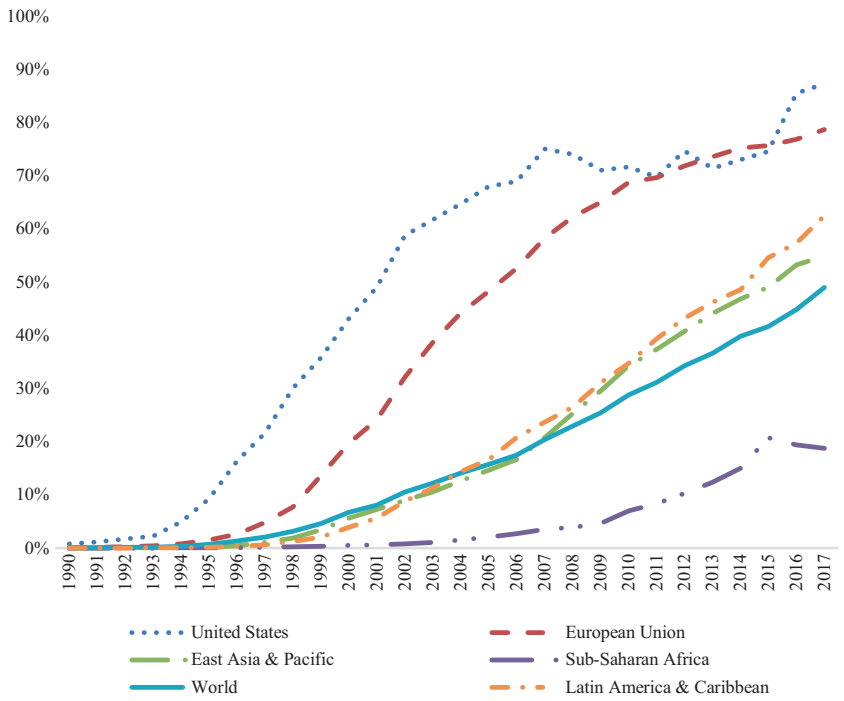

Fig. 1 Individuals using the internet as a % of population. (Source: Authors' calculations based on World Bank (2021)) data.

industrial age to the new digital era. Figure 1 shows the proportion of population using internet from 1990 up to the latest available data in different regions of the world. As it can be seen in the figure, while more than half of the population in the world uses internet, there are still significant differences according to the region considered: while in the United States or Europe more than 80% of the population are regular users, in Sub-Saharan Africa this proportion is still around 20% (although it is improving thanks to the fast penetration of smartphones). Closing the digital divide is, in fact, a policy priority in order to not leave nobody behind in the new technological world (World Bank, 2016).

But the fast developments in information and communication technologies have not only changed the way we live, and we work, it has also affected how we produce and how we consume. The improvements in

global transportation infrastructures and the associated reduction in transportation costs together with the new role played by information technologies are key factors explaining the rise of Global Value Chains (GVCs). In fact, globalisation is closely intertwined with technological change (Rodrik, 2018). Since 1991, trade openness, internationalisation of production and labour mobility have fostered productivity and innovation around the globe, increasing technological imports and technology adoption in both developed and developing countries, but also changing the nature of work.

2 Technology and Jobs

Although in the past there have been movements protesting against technological improvements (such as Ludists), there is a wide consensus among economists that technological progress during the nineteenth and the twentieth century did not result in net job losses (Autor, 2015). In fact, previous industrial revolutions seem to have generated new opportunities for low-wage workers to move into better-paid medium-skilled jobs allowing them to capture the benefits of specialisation and the economies of scale derived from automation thanks to mass education. Skill-biased technological change (SBTC) and capital-skill complementary (CSD) hypothesis explain how technological changes translated into higher economic growth and a lower overall wage inequality during most of the twentieth century (Katz & Murphy, 1992; Goldin & Katz, 2008; Buyst et al., 2018).

However, starting in the 1980s, the impacts of the current technological revolution on labour markets seem to be different. Acemoglu and Autor (2011) argue that technological progress is currently replacing workers doing routine and therefore codifiable tasks. According to this hypothesis, recent technological change is "routine-biased" (RBTC) leading to job polarisation because easily codifiable routine tasks are concentrated in intermediate jobs. On the opposite, high-paid jobs that mainly involve doing non-routine cognitive tasks with a creative component and low-paid jobs involving non-routine interactive and manual tasks seem to be much more difficult to automate. In a very influential paper, Frey and Osborne (2017) analysed the probability of computerisation for 702 occupations in the US by considering these dimensions and applying this approach; they assigned a probability of 0.94 that computerisation would lead to job losses within the next decades to accountants and auditors but

only 0.37 to actors, 0.17 to firefighters or nearly 0 to dentists. From an aggregate perspective, they found that 47% of all jobs in the US were potentially automatable in a near future.

This grimmer stance has been challenged, though. Arntz et al. (2016, 2017) argue that occupation-level approaches significantly overestimate automation potentials. In particular, they argue that within an occupation, many workers specialise in tasks that cannot be easily automated, and once this is taken into account, only about 9% of jobs in the US would be at risk of automation. As highlighted by Autor (2014), this result is also related to Polanyi's paradox ("We know more than we can tell") to explain why not more jobs have been automated and well in line with the evidence obtained by Acemoglu and Restrepo (2020). Looking at US data from 1990 and 2007, their estimates suggest that an extra robot per 1000 workers would reduce the employment to population ratio by 0.18–0.34 percentage points and wages by 0.25–0.5%. Although these are sizable effects, it is a relatively small fraction of US jobs being affected by robots. However, the pace at which machines are gaining the ability to perform cognitive tasks is much faster than in the past and labour markets have polarised due to the lower demand for mid-skill workers compared to workers in the extremes of skill distribution (Gibbs, 2017).

Technological change has also contributed to the worldwide increase of non-standard forms of employment (ILO, 2016), particularly due to the emergence of digital labour platforms. According to the recent study by ILO (2021), digital labour platforms provide new work opportunities, including some groups of workers that have been marginalised in traditional labour markets, while allowing businesses to expand their activity to new customers while accessing to a more flexible workforce with varied skills. However, there are doubts about the risks that digital platforms represent to working conditions, the regularity of work and income and the lower social protection compared to traditional jobs.

In This Section: Jobs and Technological Change

The second part in this volume, "Jobs and Technological Change," studies these issues by compiling relevant publications regarding the link between technological innovation and changes in the nature of work and labour over the last three decades.

How does international migration causally relate to the technological gap? The next chapter (Chapter "The Impact of Differences in the Levels

of Technology on International Labor Migration") in this book by Galor and Stark (2022) analyses the impact of differences in the levels of technology on international labour migration. The authors analyse which migratory patterns result from technological differences. They provide a microeconomic foundation of international labour migration, modelled in a dynamic general equilibrium framework of a world characterised by technological differences. As reported by Galor and Stark (2022), in line with the prediction of static models, if international capital movements are unrestricted, labour is likely to migrate from the technologically inferior to the technologically superior country. However, the authors show that in the absence of international capital mobility, this dynamic would not be irreversible. This is because, under certain conditions of the stationary autarkic equilibrium in the technologically superior country, migration may take place from the technologically superior country to the technologically inferior country. While also other complementary factors may impact migratory decisions, Galor and Stark innovate by isolating and exploring several technology-migration connections, and by leading to future research questions about the possible interactions of the various determinants of labour migration, such as uncertainty and availability of information, and skill differentiation.

In the developing world, labour's marketplace bargaining power of workers may be enhanced or reduced by international market integration. Globalisation can upgrade labour's conditions due to increasing levels of exports and production, and consequentially of employment, but also put pressure on governments in developing countries to keep wages competitive, worsening the bargaining power of labour. In Chapter "Are Workers in the Developing World Winners or Losers in the Current Era of Globalization?", Rudra (2022) investigates whether workers in less-developed countries (LDC) are winners or losers in the current global economy. The author uses a panel data set for 59 developing countries to test the overall impact of globalisation on labour across various levels of economic development. This is done by analysing the effects on potential labour-power (PLP) trends, which is an index defined by the author for making inferences about the extent of labour's marketplace bargaining power. Rudra (2022) finds the effects of globalisation on workers seem to depend on the level of LDC economic development. While lower income LDCs are experiencing greater employment opportunities thanks to international market integration, surplus labour problems have worsen and labour solidarity has been distressed in all countries with the exception

of high-income LDCs (e.g. in countries with the "growing population of skilled labour and the relatively small pools of surplus labour"). As Rudra points out, this suggests that workers in LDCs at higher levels of economic development are the only "winners" in the current era of globalisation. Those in middle-/low-income countries, while they have better access to employment opportunities, are "losers," in the sense that PLP does not significantly increase with markets integration.

In Chapter "The Role of Labour in Capability Upgrading: The Case of Emerging Market Multinationals", Amendolagine and Rabelloti (2022) provide a complementary perspective to this analysis by considering how access to skilled labour in developed countries can help Emerging Market Multinationals (EMNEs) to improve their technical and managerial skills. In fact, EMNEs have recently started to increasingly undertake strategic asset seeking investments in developed countries (i.e. through the acquisition of technological leading companies) with the aim of acquiring knowledge and upgrading their innovative capabilities. However, the authors explain that it is not easy for these firms to take advantage of their FDIs in developed countries to improve their innovative capabilities in the host region. Only those firms with strong "absorptive capacity and high reputational status" seem to be successful in exploiting the benefits of this strategy.

To what extent technological and structural changes are impacting employment opportunities? Employment effects of innovation are studied in Chapter "Employment Effect of Innovation", by Kancs and Siliverstovs (2022). Their study provides novel evidence about the nexus between R & D-driven innovation and firm employment in OECD countries in a continuous framework. As the authors show, while innovation can increase labour demand, nonlinearities of the innovation/employment relationship can arise from many aspects. These may be the nature and the purpose of innovation, its dimension or, from the other side of the dynamics, on the sector of the firm, institutions, the structure of the workforce skills, and other complementary factors, which may contribute to diversifying employment effects at different innovation intensities. Using R & D scoreboard data, they frame and confirm the nonlinearities embedded in the innovation/employment relationship, by noting differences in the reaction of employment to the innovation activity of the firm, depending on the actual level of the R & D intensity. Kancs and Siliverstovs' (2022) results suggest the labour-saving aspect of innovation is more

marked for firms with medium-high levels of R & D intensity, while it tends to increase with the levels of R & D intensity.

Concerning the relation between biased technological change and employment risks related to premature deindustrialisation, Abbot, Tarp and Wu (2022) provide new evidence about the Vietnamese development experience in Chapter "Structural Transformation, Biased Technological Change and Employment in Vietnam". The authors analyse the differences observed between employment and GDP growth to examine the specific roles played by structural change of the economy, technical innovation and institutional bias towards capital-intensive development. The period considered by the authors is from 2000 to 2011, while the data used are from seven aggregate sectors and the overall Vietnamese economy. The study suggests only a partial portion of the difference between employment and GDP growth is driven by shifts from low-productivity to high-productivity sectors (e.g. structural transformation). Instead, a significant share of the employment-growth difference is due to technical change, when low elasticities of substitution and labour-augmenting bias in technical change are taken into account in the analysis. A remaining small portion of the difference is also explained by capital-intensive investments by state-owned enterprises. Abbot, Tarp and Wu's (2022) results are consistent with significant labour-augmenting technical progress, and better explain recent economic performance in Vietnam. They also suggest that restructuring and investments allocation into private firms could be an important factor for driving future employment growth.

Innovation may also be considered as the specific tool of entrepreneurs (Drucker, 1985) with its own economic and social dimension, and not only a simplistic technical change. Nonetheless, innovation and entrepreneurship are too often imagined at the top-tier countries and individuals, neglecting other realities and rationalities. The second part of the book continues with Chapter "'Helping a Large Number of People Become a Little Less Poor': The Logic of Survival Entrepreneurs", by Berner, Gomez and Knorringa (2022), where the survival and growth-oriented entrepreneurs' typologies are studied. The authors outline and systematise the different characteristics of the various categories of entrepreneurs to better understand their behaviour and different needs. In the chapter, Berner, Gomez and Knorringa systematically develop the policy implications for supporting survival entrepreneurs. The authors suggest there is a qualitative difference between survival and growth-oriented enterprises concerning several key variables, and this is reflected in

different typologies and rationales of entrepreneurship. They stress that great part of the existing policy interventions are based on the idea that all entrepreneurs are growth-oriented, which aim to push "entrepreneurs onto the ladder of graduation towards more specialization and growth orientation." This is unlikely to address the specific needs of survivalists, who are usually more concerned with poverty, structural uncertainties and access to basic services and infrastructure.

The inclusiveness of innovation processes is still debated. Given the lack of capital for more expensive technologies, frugal innovation has become increasingly important merging different pieces of literature (Leliveld & Knorringa, 2018). While innovative systems or services can promote employment opportunities and entrepreneurship, they may also contribute to informalisation, fostering inequalities and capitalist exploitation. Chapter "The 'Transformational' Potential of Mobile Money in Zambia: Mobile Money Agents as Drivers of Entrepreneurship in Kitwe?" in this book, by Peša (2022), surveys this issue by analysing whether the service of mobile money can become a pro-poor frugal innovation for the agents and tellers in Kitwe, Zambia. The author proposes an empirical study focusing on entrepreneurship and labour relations, which suggests the existence of differences between agents and tellers in mobile money services. Peša (2022) studies these differences by focusing on the educational and employment background of agents and tellers, financing and business operations and the franchising structure of the sector in Zambia. The chapter suggests that frugal innovation studies should pay more attention to the longer term and ambiguous outcomes which frugal innovations have on livelihoods and development. In Zambia, while processes of inclusion might be selective, mobile money might give rise to new and partially unpredictable inequalities among the various actors involved in the process. Peša proposes frugal innovation cannot be analysed in terms of either development or exploitation, and that the role of policy in frugal innovation should receive more consideration. In Chapter "Dynamics of Mobile Money Entrepreneurship and Employment in Kitwe, Zambia", Kabala (2022) updates this research by providing a new assessment of the capacity of mobile money agents to enhance financial inclusion and create employment opportunities. The conclusions obtained from interviews with agents and tellers of the three main players in the mobile money services market in Kitwe, Zambia, confirm the results by Peša (2022): only firms that can offer innovation in terms of a broad line of diverse and affordable services are the ones that survive the fierce competition, but even in the case of

successful firms, this does not avoid the existence of vast inequalities between agents and tellers.

References

Abbot, P., Tarp, F., & Wu, C. (2022). Structural Transformation, Biased Technological Change and Employment in Vietnam. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Acemoglu, D., & Autor, D. (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings," Handbook of Labor Economics. In O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics* (edition 1, Vol. 4, Chap. 12, pp. 1043–1171). Elsevier.

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy, 128*(6), 2188–2244.

Amendolagine, V., & Rabellotti, R. (2022). The Role of Labour in Capability Upgrading: The Case of Emerging Market Multinationals. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Arntz, M., Gregory, T., & Zierahn, U. (2016). *The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis*. OECD Social, Employment and Migration Working Papers, 189

Arntz, M., Gregory, T., & Zierahn, U. (2017). Revisiting the Risk of Automation. *Economics Letters, 159*, 157–160.

Autor, D. (2014). *Polanyi's Paradox and the Shape of Employment Growth*. NBER WP 20485

Autor, D. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives, 29*(3), 3–30.

Berner, E., Gomez, G., & Knorringa, P. (2022). Helping a Large Number of People Become a Little Less Poor': The Logic of Survival Entrepreneurs. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Buyst, E., Goos, M., & Salomons, A. (2018). Job Polarization: An Historical Perspective. *Oxford Review of Economic Policy, 34*(3), 461–474.

Drucker, P. (1985). *Innovation and Entrepreneurship*. Butterworth-Heinemann, Oxford.

Frey, C. B., & Osborne, M. A. (2017). "The future of employment: How susceptible are jobs to computerisation?", *Technological Forecasting and Social Change* (Vol. 114, pp. 254–280). https://doi.org/10.1016/j.techfore.2016.08.019

Galor, O., & Stark, O. (2022). The Impact of Differences in the Levels of Technology on International Labor Migration. In P. Goulart, R. Ramos, &

G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Gibbs, M. (2017). How is new technology changing job design? *IZA World of Labor* 2017: 344. https://doi.org/10.15185/izawol.344

Goldin, C., & Katz, L. (2008). *The Race Between Education and Technology*. Belknap Press for Harvard University Press.

ILO (2016). *World Employment and Social Outlook: Trends 2016*. ILO, Geneva.

ILO (2021). *World Employment and Social Outlook 2021: The role of digital labour platforms in transforming the world of work*. ILO, Geneva.

Kabala, E. (2022). Dynamics of Mobile Money Entrepreneurship and Employment in Kitwe, Zambia. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Kancs, A., & Siliverstovs, B. (2022). Employment Effect of Innovation. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Katz, L. F., & Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics, 107*(1), 35–78. https://doi.org/10.2307/2118323

Leliveld, A., & Knorringa, P. (2018). Frugal Innovation and Development Research. *European Journal of Development Research, 30*, 1–16.

Peša, I. (2022). The Developmental Potential of Frugal Innovation among Mobile Money Agents in Kitwe, Zambia. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Rodrik, D. (2018). *New Technologies, Global Value Chains, and the Developing Economies, Pathways for Prosperity Commission*. University of Oxford.

Rudra, N. (2022). Are Workers in the Developing World Winners or Losers in the Current Era of Globalization? In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

World Bank. (2016). *World Development Report 2016: Digital Dividends*. World Bank Group.

World Bank. (2021). *World Bank Open Data*. https://data.worldbank.org

The Impact of Differences in the Levels of Technology on International Labor Migration

Oded Galor and Oded Stark

1 Introduction

Among other things, countries differ in their stock of technological knowledge. Quite often, the technological stock is country specific and remains so even though production factors such as labor and capital are mobile across countries. Relatively little effort has been made either to examine the extent to which given differences in technology entail labor migration or to establish conditions under which a labor migration process will continue or come to a halt. Since certain countries appear to consistently

The original version of the chapter has been revised. A correction to this chapter can be found at https://doi.org/10.1007/978-3-030-89258-6_29

This chapter, presented at the 6th World Congress of the Econometric Society, Barcelona, August 22–29, 1990, is a revised version of Discussion Paper No. 34, Harvard University Migration and Development Program.

O. Galor (✉)
Brown University, Providence, RI, USA
e-mail: oded_galor@brown.edu

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_12

maintain technological superiority over other countries, we find it natural to pose and address the question: how does international migration causally relate to the technological gap?

In this chapter we take the technology gap *as given* and investigate its migration repercussions. Explaining the formation and persistence of a gap is outside the scope of the current chapter, but a brief reference is in order. By technology we mean more than the mechanical process of turning iron into steel; we mean useful knowledge and experience, institutions and organizational forms (such as the modern corporation and stock exchanges), and even norms and values (such as the work ethic and property rights) that impinge upon and govern the processes of production and exchange. Although countries differ in their historical heritage—size, resources, institutions—a difference which could plausibly give rise to an "initial" technological gap, why is it that over time the gap does not close? One reason may have to do with the link between the evolution of technology and actual production. Inter alia, the former depends on there being an active search for new techniques—new ways of doing things *while* doing things. To the extent that a difference in initial endowments translates into a difference in the levels of production, the additions to the existing stocks of technology will also differ and hence the technology gap will not close. Furthermore, even though countries can partake in the portion of the stock of knowledge which is common, especially through trade and international relations, *when* and *what* they take is a selection process that depends on factors and characteristics which, in themselves, differ among countries. This difference arises partly due to the historical heritage and partly due to ongoing decisions pertaining to investments in higher education, the proportion of the national human capital devoted to acquisition of new knowledge, and so on. To our judgment, just as individuals within (as well as across) countries differ in their productive skills—partly due to a difference in endowments and partly due to different histories of investment decisions in skill acquisition—so do nations differ in their "skills." And just as it is important to recognize that considerable inter-personal skill differences have a tendency to persist and hence their

O. Stark (✉)
University of Bonn, Bonn, Germany

University of Warsaw, Warsaw, Poland
e-mail: ostark@uni-bonn.de

repercussion for the inter-personal flows of income ought to be investigated, so it is appropriate to study the impact of a (given) technology gap between countries on the inter-country distribution of labor, that is, on international migration. Somewhat surprisingly, hitherto this line of inquiry has not been pursued systematically. Hence, the current chapter.

We start with quite simple considerations. Assume a static, perfectly competitive two-country world wherein capital and labor are combined in the production of a single good. The two countries are identical except for one attribute: their production technologies differ; one country exhibits a Hicks-neutral technological superiority.[1] It follows immediately that if international capital mobility is not allowed yet labor can move freely, labor will migrate from the technologically inferior to the technologically superior country. Migration will entail equalization of the wage rates across countries at which point it will cease. Yet the rental rate will not equalize; the rental rate in the technologically superior country will remain higher than the rental rate in the technologically inferior country. If, however, capital mobility is allowed as well, both labor and capital will move from the technologically inferior to the technologically superior country. Indeed, in this case all production factors will concentrate in the technologically superior country.[2]

When international capital mobility is prohibited, the patterns of labor migration in the described static world depend on two critical considerations: whether the capital-labor ratios are identical across countries and whether labor migration is motivated exclusively by (inter-country) wage differentials. When international capital mobility is unrestricted, the patterns of migration depend solely on differences in the inter-country returns to labor.

Are the predictions of the static model concerning the patterns of migration in a world that is characterized by international differences in technology indeed plausible? If, in the short run, the capital-labor ratios across countries are identical, the returns to the factors of production are higher in the technologically superior country. As this affects the patterns of savings and capital formation, differences in the capital-labor ratios are bound to arise in the long run. This in turn will impinge upon migration

[1] The role of international differences in technology in international trade theory has been analyzed by Findlay and Gubert (1959), Bhagwati (1964), Ruffin (1984), as well as others.

[2] Bhagwati and Srinivasan (1983) address the choice between capital and labor movement. Wong (1983) studies the choice between trade in goods and factor movements.

incentives. Furthermore, as will become evident from the dynamic model presented in this chapter, inter-country wage differentials do not constitute all the considerations of potential migrants.

In the current chapter we analyze the pattern of international labor migration in a world characterized by international differences in technologies and by identical individuals within as well as across countries. The analysis is conducted within a dynamic general equilibrium framework for a two-country, one-good two-factor perfectly competitive world where each country is characterized by an overlapping-generations model along the lines of Diamond (1965). This framework allows us to trace the dynamic considerations of potential migrants as a natural consequence of intertemporal utility maximization, and to explicitly incorporate into the analysis the impacts of technological differences on both capital formation and factor returns.

In the absence of international capital mobility, the analysis indicates that in contrast to the prediction of the static models, labor may migrate from the technologically superior country to the technologically inferior country. This occurs if and only if in the technologically superior country the stationary autarkic equilibrium is characterized by over-investment relative to the Golden Rule and the long-run elasticity of the stationary autarkic interest rate with respect to the technological level is sufficiently large. However, if international capital movements are unrestricted, the prediction of this dynamic model tallies with that of the static models: migration takes place, unconditionally, from the technologically inferior country to the technologically superior country.

Two clarifications are in order. First, abstracting from international differences in time preferences within as well as across countries results in a somewhat unusual view of migration. Migration plays an equilibrating role solely in the short run. In the long run, however, it plays neither an equilibrating nor a disequilibrating role. Migration results in a concentration of population in the country with the higher stationary lifetime utility and non-migrants' stationary welfare is unaffected. The long-run neutrality of migration stems in part from the assumption that individuals are identical within as well as across countries. Migrants perfectly assimilate into the recipient country and consequently do not affect the population composition in either country. The patterns and the welfare implications of international labor migration in a two-country overlapping-generations world in which individuals differ within as well as across countries in their time preferences, but technologies are identical, are analyzed in Galor (1986), where

the long-run neutrality of migration is naturally eliminated. Migration plays an equilibrating role affecting non-migrants' stationary welfare.

Second, as pointed out at the outset, the current chapter does not attempt to explain the technology gap across countries; it is taken as exogenously given. The proposed model and its predictions would be subsequently weakened if the technological differences across countries are systematically linked with other potential differences (e.g., time preferences) across countries. Further analyses, based on endogenous technological differences, are therefore clearly desirable.

2 The World Economy in the Absence of International Factor Movements

Consider an infinite time horizon world where economic activity is performed under perfect competition and certainty. The world consists of two countries. At any period of time capital and labor are combined in the production of a single good. In the absence of international factor movements, L_t^i, the endowment of labor in country i, $i = A, B$, at time t, is exogenously given, whereas K_t^i, the endowment of capital in country i at time t, is the output produced but not consumed in the preceding period in country i:

$$L_t^i = L^{-i}; \tag{1}$$

$$K_t^i = Y_{t-1}^i - C_{t-1}^i, \tag{2}$$

where Y_{t-1}^i and C_{t-1}^i are respectively the aggregate production and consumption at time $t - 1$ in country i. Thus, capital is fully depreciated after a single period and the population does not grow.[3]

Production

Production occurs within a period according to a constant-returns-to-scale production function. The function is invariant through time. The output produced in country i at time t, Y_t^i, is

[3] These assumptions are chosen in order to simplify the exposition. The analysis could have been conducted under any feasible rates of population growth and capital depreciation with the main results remaining intact.

$$Y_t^i = \alpha^i F\left(K_t^i, L_t^i\right) = \alpha^i L_t^i f\left(k_t^i\right); \quad k_t^i = K_t^i / L_t^i, \tag{3}$$

where $\alpha^i > 0$ is the technological coefficient in country i. For $\varepsilon > 0$ which is sufficiently small, $\alpha^i + \varepsilon = \alpha^j$, $i, j = A, B$; $i \neq j$. Namely, country j has an infinitesimal Hicks-neutral technological superiority over country i.

The production function is twice continuously differentiable, strictly monotonic increasing and concave, and is defined on the input space R_+^2:

$$\begin{aligned} &f'\left(k^i\right) > 0, \quad f''\left(k^i\right) \left\langle 0, \quad f\left(k^i\right) \right\rangle 0, \quad \forall k^i > 0; \\ &\lim_{k^i \to 0} f\left(k^i\right) = \lim_{k^i \to \infty} f'\left(k^i\right) = 0; \quad \lim_{k^i \to \infty} f'\left(k^i\right) = \infty. \end{aligned} \tag{4}$$

The producers operate in a perfectly competitive environment. The inverse demands for factors of production are therefore characterized by the first-order conditions for profit maximization:

$$w_t^i = \alpha^i \left[f\left(k_t^i\right) - f'\left(k_t^i\right) k_t^i \right]; \tag{5}$$

$$r_t^i = \alpha^i f'\left(k_t^i\right), \tag{6}$$

where w_t^i and r_t^i are respectively the wage and the rental rate at time t in country i; output is the numeraire.

Consumption and Factor Supply

In every period t, L^{-i} individuals are born in country i. Individuals are identical within as well as across time. They live two periods. In the first period they work and earn the competitive market wage, w_t^i, and in the second they are retired. Individuals born at time t in country i are characterized by their intertemporal utility function $u\left(c_t^{t,i}, c_{t+1}^{t,i}\right)$ defined over consumption during the first and second periods of their lives, as well as by their unit-endowment of labor during the first period of their lives.[4]

The intertemporal utility function is twice continuously differentiable, strictly monotonic increasing, strictly quasi-concave and is defined on the consumption set R_+^2. For every $c_t > 0$ and $c_{t+1} > 0$

[4] Intergenerational altruism is not considered. Kemp and Kondo (1986) incorporate intergenerational altruism endogenizing the population growth modeled in Galor (1986).

$$u_1\left(0,c_{t+1}^t\right)=u_2\left(c_t^t,0\right)=\infty;\quad u_1\left(\infty,c_{t+1}^t\right)=u_2\left(c_t^t,\infty\right)=0. \tag{7}$$

At time t young individuals in country i supply their unit-endowment of labor inelastically and allocate the resulting income, w_t^i, between first-period consumption, $c_t^{t,i}$, and savings, $s_t^{t,i}$, so as to maximize their utility function $u\left(c_t^{t,i},c_{t+1}^{t,i}\right)$. The individual's budget constraint is therefore $c_t^{t,i}+s_s^{t,i}\le w_t^i$. Since capital is the only store of value in each country and international capital movements are not permitted, second-period consumption is $c_{t+1}^{t,i}=r_{t+1}^i s_t^{t,i}$.

Thus, the maximization problem faced by a young individual in country i at time t is

$$\max u\left(c_t^{t,i},c_{t+1}^{t,i}\right)$$

subject to

$$c_t^{t,i}+\left[1/r_{t+1}^i\right]c_{t+1}^{t,i}\le w_t^i;\quad c_t^{t,i}\ge 0,\quad c_{t+1}^{t,i}\ge 0. \tag{8}$$

Given the properties of the utility function, a solution to the intertemporal optimization exists for strictly positive prices and is unique. The optimal consumption vector of an individual of generation t in country i is $\left[c_t^t\left(w_t^i,r_{t+1}^i\right),c_{t+1}^t\left(w_t^i,r_{t+1}^i\right)\right]$. The individual's savings implied by the solution to (8) are

$$s_t^{t,i}=s\left(w_t^i,r_{t+1}^i\right)=w_t^i-c_t^t\left(w_t^i,r_{t+1}^i\right). \tag{9}$$

It is assumed that $\partial s_t/\partial r_{t+1}\ge 0$ and $\partial s_t/\partial w_t>0$ (i.e., savings are a non-decreasing function of the interest rate and second-period consumption is a normal good).

A Stationary Equilibrium

Definition. In the absence of international movements of factors of production an autarkic stationary equilibrium in country i is a stationary price sequence $\left\{\hat{w}^i,\hat{r}^i\right\}$ under which in every period the demand for labor in country i equals its aggregate supply, L^{-i}, and the demand for capital equals the aggregate supply of savings in the country, $L^{-i}s\left(\hat{w}^i,\hat{r}^i\right)$. Namely,

$$\hat{w}^i = \alpha^i \left\{ f\left[s\left(\hat{w}^i, \hat{r}^i\right)\right] - f'\left[s\left(\hat{w}^i, \hat{r}^i\right)\right] s\left(\hat{w}^i, \hat{r}^i\right)\right\}; \tag{10}$$

$$\hat{r}^i = \alpha^i f'\left[s\left(\hat{w}^i, \hat{r}^i\right)\right]. \tag{11}$$

Remark 1: Under stationary equilibrium the level of utility attained by an individual in country i is

$$v\left(\hat{w}^i, \hat{r}^i\right) = u\left[c_1\left(\hat{w}^i, \hat{r}^i\right), c_2\left(\hat{w}^i, \hat{r}^i\right)\right]. \tag{12}$$

Lemma 1: *The steady-state equilibrium in country i displays local stability if*

$$1 - \alpha^i f''\left(\hat{s}^i\right)\left[\frac{\partial \hat{s}^i}{\partial r^i} - \hat{s}^i \frac{\partial \hat{s}^i}{\partial w^i}\right] > 0, \tag{13}$$

where $\hat{s}^i = s\left(\hat{w}^i, \hat{r}^i\right)$.

Proof: The dynamical system is characterized by (5), (6), and

$$k^i_{t+1} = s_t\left(w^i_t, r^i_{t+1}\right). \tag{14}$$

Local stability is satisfied if

$$\left|\frac{\mathrm{d}k^i_{t+1}}{\mathrm{d}k^i_t}\right|_{k^i = \hat{k}^i} < 1. \tag{15}$$

Consequently, local stability is satisfied if

$$\left|\frac{-\alpha^i f''\left(\hat{s}^i\right)\hat{s}^i\left(\partial \hat{s}^i / \partial w^i\right)}{1 - \alpha^i f''\left(\hat{s}^i\right)\left(\partial \hat{s}^i / \partial r^i\right)}\right| < 1. \tag{16}$$

Noting that, by assumption, $\partial \hat{s}^i / \partial w^i > 0$ and $\left(\partial \hat{s}^i / \partial r^i\right) \geq 0$, the lemma follows immediately from (16).

Remark 2: The sole distinction between the two countries is the Hicks-neutral difference in technology.

3 The Pattern of International Labor Migration

International Capital Mobility Is Not Permitted

Suppose that the sufficient conditions for the existence of a unique and globally stable nontrivial steady-state equilibrium (Galor and Ryder (1989)) are satisfied and that each country is at its stationary autarkic equilibrium. International labor migration is now allowed, yet international capital mobility is not permitted. Individuals spend their entire lifetime in either the home country or the receiving country (in which case, to recall, with regard to their work, consumption, and savings behavior, they become perfect replicas of the individuals in the country they join). Thus, incentives for labor migration from country i to country j exist if and only if

$$v\left(\hat{w}^j,\hat{r}^j\right) > v\left(\hat{w}^i,\hat{r}^i\right) \quad i,j = A,B; \quad i \neq j. \tag{17}$$

Namely, incentives exist if and only if the utility level attained by individuals who migrate to country j and face its stationary autarkic equilibrium prices $\left(\hat{w}^j,\hat{r}^j\right)$ is higher than that attained in the country of origin where individuals face the stationary autarkic prices, $\left(\hat{w}^i,\hat{r}^i\right)$.[5]

Remark 3: Whereas in a static one-good two-factor world incentives for international labor migration are determined solely by the international differences in wages, in the current dynamic framework the incentives for migration are determined by international differences in the stationary autarkic indirect utilities, which in turn reflect the differences in the stationary autarkic baskets of wages and interest rate weighted according to the intertemporal preferences.

Suppose that migrants' children are born in the receiving country. Namely, the labor endowment at time $t+1$ in country i, L^i_{t+1}, is

$$L^i_{t+1} = L^i_t + m^i_t \tag{18}$$

[5] Since time preferences are identical across countries, the phenomenon of bilateral migration as presented in Galor (1986) cannot occur. Furthermore, we do not associate migration with a change in longevity. Clearly, if migration is associated with a change in lifespans, savings and thus capital-labor ratios will be affected. Consequently, migration patterns and welfare will be affected as well.

where m_t^i is the net inflow of migration to country i, at time t.

Remark 4: Since individuals (and thus time preferences) are identical within as well as across countries, migration has no effect on the stationary savings per capita, the stationary capital-labor ratio, and thereby on the stationary autarkic equilibrium prices.

Lemma 2: *The actual patterns of labor migration during the transition period from the pre-migration to the post-migration stationary equilibrium follow the direction determined by the initial incentives for migration given by condition (17).*

Proof: Suppose, without loss of generality, that $v\left(\hat{w}^B,\hat{r}^B\right) > v\left(\hat{w}^A,\hat{r}^A\right)$. Namely, incentives exist for migration from country A to country B. Suppose that at time t_0 (within the stationary state) labor is permitted to migrate internationally. Young individuals in country A at time t_0 will find it beneficial to migrate to country B. The number of individuals m_0, who will actually migrate from country A to country B at time t_0, will be determined by the speed of adjustment in the capital-labor ratios and factor prices in the two countries. The capital-labor ratio in country B will be reduced from $\hat{k}^B$ to k_0^B, whereas the capital-labor ratio in country A, will rise from $\hat{k}^A$ to k_0^B. The associated wages w_0^A and w_0^B, and the expected interest rates r_1^A and r_1^B are such that $v\left(w^B,r^B\right) = v\left(w_0^A,r_1^A\right)$.[6] Namely, under these prices there are no further incentives for migration in period t_0. In period $t_1 \equiv t_0 + 1$, as follows from (18), more people are born in country B and less in A relative to period t_0. In the absence of further migration, the capital-labor ratios in the two countries will converge to their autarkic steady-state levels. (Note that the stationary equilibrium is globally stable.) Thus, incentives for migration from A to B will ultimately be restored. The process will continue till all individuals migrate from A to B.[7]

Proposition 1: *Consider a stable stationary autarkic equilibrium of a two-country overlapping-generations world where production technologies differ across countries. If restricted or unrestricted international labor migration is permitted, whereas international capital mobility is not allowed,*

[6] If equalization does not occur, then the lemma is satisfied trivially (i.e., following the initial incentives for migration all individuals from country ***A*** migrate to country ***B***).

[7] This prediction cannot be brushed aside on the basis of the argument that in reality countries never lose their entire labor force through migration. In reality labor migration *is* restricted.

(a) *labor migrates from the technologically superior country (to the technologically inferior country) if and only if*

(i) *the country's stationary autarkic equilibrium is characterized by over-investment relative to the Golden-Rule, that is,* $\hat{r}^i < 1$, *and*

(ii) *the country's long-run elasticity of the stationary autarkic interest rate with respect to the technological level is sufficiently large, that is,* $\hat{\eta}^i_{r,\alpha} > \alpha^i f\left(\hat{s}^i\right) / \hat{s}^i \left(1 - \hat{r}^i\right)$.

(b) *labor migrates from the technologically inferior country (to the technologically superior country) otherwise.*

Proof: As was established in Lemma 2, international labor migration is in the direction determined by the initial migration condition, which is to say that the pattern of migration is determined by the international differences in the utility levels attained under the stationary autarkic equilibrium. Thus, labor migrates from the technologically superior country, *j*, if and only if $v\left(\hat{w}^i, \hat{r}^i\right) > v\left(\hat{w}^j, \hat{r}^j\right)$, whereas labor migrates from the technologically inferior country, *i*, if and only if $v\left(\hat{w}^i, \hat{r}^i\right) < v\left(\hat{w}^j, \hat{r}^j\right)$. Noting that the sole difference between the two countries is the technological level, it follows that $v\left(\hat{w}^i, \hat{r}^i\right) \gtrless v\left(\hat{w}^j, \hat{r}^j\right)$ if and only if $\mathrm{d}v\left(\hat{w}^i, \hat{r}^i\right) / \mathrm{d}\alpha^i \gtrless 0$, where, to recall, α^i is the technological parameter in country *i*, $\alpha^i + \varepsilon = \alpha^j$ and $\varepsilon > 0$ is sufficiently small.

Totally differentiating the indirect utility function at the stationary autarkic equilibrium, it follows that

$$\mathrm{d}v\left(\hat{w}^i, \hat{r}^i\right) = \frac{\partial v\left(\hat{w}^i, \hat{r}^i\right)}{\partial w^i} \mathrm{d}\hat{w}^i + \frac{\partial v\left(\hat{w}^i, \hat{r}^i\right)}{\partial r^i} \mathrm{d}\hat{r}^i \tag{19}$$

Using the envelope theorem

$$\frac{1}{\hat{\lambda}^i} \mathrm{d}v\left(\hat{w}^i, \hat{r}^i\right) = \left[1, \hat{s}^i / \hat{r}^i\right] \mathrm{d}\hat{p}^i, \tag{20}$$

where $\mathrm{d}\hat{p}^i \equiv \left[\mathrm{d}\hat{w}^i, \mathrm{d}\hat{r}^i\right]$, and $\lambda^i \equiv \partial v\left(\hat{w}^i, \hat{r}^i\right) / \partial w^i$ is the marginal utility of income. $\left(1 / \hat{\lambda}^i\right) d\hat{v}^i$ is, therefore, the total change in the stationary per-capita real income in country *i*.

Using (10) and (11) it follows that

$$\hat{w}^i = \alpha^i f\left(\hat{s}^i\right) - \hat{r}^i \hat{s}^i. \tag{21}$$

Thus, noting that (11) implies that $\left[\alpha^i f'\left(\hat{s}^i\right) - \hat{r}^i\right]\partial s^i / \partial \alpha^i = 0$,

$$\frac{\mathrm{d}\hat{w}^i}{\mathrm{d}\alpha^i} = f\left(\hat{s}^i\right) - \hat{s}^i \frac{\mathrm{d}\hat{r}^i}{\mathrm{d}\alpha^i}. \tag{22}$$

Substituting into (20)

$$\frac{1}{\hat{\lambda}^i}\frac{\mathrm{d}v\left(\hat{w}^i, \hat{r}^i\right)}{\mathrm{d}\alpha^i} = f\left(\hat{s}^i\right) + \hat{s}^i \frac{\left(1-\hat{r}^i\right)}{\hat{r}^i}\frac{\mathrm{d}\hat{r}^i}{\mathrm{d}\alpha^i}. \tag{23}$$

Rearranging terms, it follows that

$$\frac{1}{\hat{\lambda}^i}\frac{\mathrm{d}v\left(\hat{w}^i, \hat{r}^i\right)}{\mathrm{d}\alpha^i} = \frac{\left(1-\hat{r}^i\right)}{\hat{r}^i} f\left(\hat{s}^i\right)\left[\frac{\hat{r}^i}{\left(1-\hat{r}^i\right)} - \hat{\xi}^i \hat{\eta}^i_{r,\alpha}\right], \tag{24}$$

where $\hat{\xi}^i \equiv \hat{r}^i \hat{s}^i / \alpha^i f\left(\hat{s}^i\right)$ is the share of capital in total output in country i, and $\hat{\eta}^i_{r,\alpha} \equiv -\left[\mathrm{d}\hat{r}^i / \mathrm{d}\alpha^i\right]\left[\alpha^i / \hat{r}^i\right]$ is the long-run elasticity of the stationary autarkic interest rate with respect to the technological level.

Lemma 3:

$$\frac{\mathrm{d}\hat{w}^i}{\mathrm{d}\alpha^i} = \frac{1}{\alpha^{i\#} D^{i\#}}\left[\hat{w}^i - \left(\alpha^i\right)^2 f''\left(\hat{s}^i\right) f\left(\hat{s}^i\right)\frac{\partial \hat{s}^i}{\partial r^i}\right] > 0 \tag{25}$$

$$\frac{\mathrm{d}\hat{r}^i}{\mathrm{d}\alpha^i} = \frac{1}{\alpha^{i\#} D^{i\#}}\left[\hat{r}^i + \left(\alpha^i\right)^2 f''\left(\hat{s}^i\right) f\left(\hat{s}^i\right)\frac{\partial \hat{s}^i}{\partial w^i}\right] \begin{matrix} < \\ = \\ > \end{matrix} 0. \tag{26}$$

Proof: Totally differentiating the stationary equilibrium conditions (10) and (11) it follows that

$$\mathrm{d}\hat{p}^i = \alpha^i M^i \mathrm{d}\hat{s}^i + \left(\hat{p}^i / \alpha^i\right)\mathrm{d}\alpha^i, \tag{27}$$

where

$$\mathrm{d}\hat{s}^i = \left[\partial\hat{s}^i / \partial w^i, \partial\hat{s}^i / \partial r^i\right]\mathrm{d}\hat{p}^i \tag{28}$$

and

$$M \equiv \left[-f''\left(\hat{s}^i\right)\hat{s}^i, f''\left(\hat{s}^i\right)\right]' \quad \text{and} \quad \mathrm{d}\hat{p}^i = \left[\mathrm{d}\hat{w}^i, \mathrm{d}r^i\right]'. \tag{29}$$

Thus,

$$\mathrm{d}\hat{p}^i = \left[I - \alpha^i M^i\left(\partial\hat{s}^i / \partial p^i\right)\right]^{-1}\left(\hat{p}^i / \alpha^i\right)\mathrm{d}\alpha^i, \tag{30}$$

where I is the 2×2 identity matrix.

Let $D^i \equiv \left[I - \alpha^i M^i\left(\partial\hat{s}^i / \partial p^i\right)\right]$. Then it follows that

$$d\hat{p}^i = \frac{1}{\alpha^{i\#}\, D^{i\#}}\mathrm{Adj}D^i\,\hat{p}^i\mathrm{d}\alpha^i, \tag{31}$$

where $|D^i|$ is the determinant of D^i. If the stationary equilibrium is locally stable then $|D^i| > 0$ as established in Lemma 1. Adj D^i is the transpose of the matrix of the co-factors of D^i,

$$\mathrm{Adj}D^i = \begin{bmatrix} 1 - \alpha^i f''\left(\hat{s}^i\right)\dfrac{\partial\hat{s}^i}{\partial r} & -\alpha^i f''\left(\hat{s}^i\right)\hat{s}^i\dfrac{\partial\hat{s}^i}{\partial r} \\ \alpha^i f''\left(\hat{s}^i\right)\dfrac{\partial\hat{s}^i}{\partial w} & 1 + \alpha^i f''\left(\hat{s}^i\right)\hat{s}^i\dfrac{\partial\hat{s}^i}{\partial w} \end{bmatrix}. \tag{32}$$

Upon substitution into (30), noting that $\alpha^i f\left(\hat{s}^i\right) = \hat{w}^i + \hat{r}^i\hat{s}^i$, the lemma follows. □

Lemma 4: *Migration from the technologically superior country to the technologically inferior country occurs only if* $\mathrm{d}r^i/\mathrm{d}\alpha^i < 0$.

Proof: Using (20)

$$\frac{1}{\hat{\lambda}^i}\frac{\mathrm{d}v\left(\hat{w}^i,\hat{r}^i\right)}{\mathrm{d}\alpha^i}=\frac{\mathrm{d}\hat{w}^i}{\mathrm{d}\alpha^i}+\frac{\hat{s}^i\mathrm{d}\hat{r}^i}{\hat{r}^i\mathrm{d}\alpha^i}. \tag{33}$$

Thus, since $\left[\mathrm{d}\hat{w}^i/\mathrm{d}\alpha^i\right]>0$ as established in Lemma 2, the lemma follows. □

Applying Lemma 3 and Lemma 4 to (24), the proposition follows, noting that $r = 1$ if the stationary equilibrium coincides with the Golden Rule. (Recall that full depreciation of capital and zero population growth were assumed.) □

Interpretations: Proposition 1 demonstrates that labor migrates from the technologically superior country to the technologically inferior country if two conditions are fulfilled. The first is that the stationary autarkic equilibrium of the technologically inferior country is characterized by over-investment relative to the Golden Rule. The second is that the long-run elasticity of the stationary autarkic interest rate with respect to the technological parameter must be sufficiently large.

A necessary condition for migration from the technologically superior country is, therefore, a significantly higher rate of interest in the technologically inferior country. Satisfaction of this requirement is demonstrated to be compatible with the local stability of the stationary equilibrium. If the interest rate is indeed higher in the technologically inferior country, then this country has a stationary capital-labor ratio and, thus, stationary output per capita which are lower than those of the technologically superior country.

Consequently, migration from the technologically superior country involves a movement to a country in which the output which is available for distribution between young and old at any point in time is lower. Since, however, the market distribution in the technologically superior country is characterized by dynamic inefficiency (manifested by over-investment relative to the Golden Rule) migration to the technologically inferior country where the interest rate is higher and, thus, closer to the Golden Rule represents an improvement in the intertemporal allocation of wages between first- and second-period consumption. Under the conditions specified in Proposition 1 the improvement in intertemporal efficiency dominates the loss of output from the viewpoint of a migrant (i.e., the improvement in the interest rate is evaluated more than the deterioration in the wage rate). Thus, labor migrates from the technologically superior country to the technologically inferior country.

Unrestricted International Capital Mobility

Suppose that at the stationary autarkic equilibrium unrestricted international capital movements are permitted. A new stationary equilibrium is achieved in each of the countries, where in the technologically inferior country i,

$$\tilde{r}^{i} = \tilde{r}^{j}, \tag{34}$$

$$\tilde{w}^{i} < \tilde{w}^{j}. \tag{35}$$

Proposition 2: *Consider a stable stationary equilibrium of a two-country overlapping-generations world where technologies differ across countries and international capital mobility is unrestricted. If international labor migration, restricted or unrestricted, is permitted, labor migrates from the technologically inferior to the technologically superior country.*

Proof: Noting (17), since $r^{i} = r^{j}$, an incentive for migration from country i to country j exists if $w^{j} > w^{i}$. Given the fact that capital movements are unrestricted, $w^{j} > w^{i}$ (where j is the technologically superior country) and hence, incentives exist for migration from country i to country j. The pattern of migration follows the initial incentive since, as long as there is labor in country i, $r^{i} = r^{j}$ and thus, given the technological superiority of country j, $w^{j} > w^{i}$.

4 Conclusions

In this chapter we provide a microeconomic foundation of international labor migration modeled in a general equilibrium framework of a world characterized by technological differences. Specifically, we have asked which migratory patterns result from technological differences. Although our analysis has utilized a number of specific assumptions, our main results are quite robust to their relaxation. For example, characterizing the technological difference as capital augmenting or labor augmenting will not alter the basic migration rule as stated in Proposition 1; it will only affect the critical magnitude of the long-run elasticity of the interest rate with respect to the technological parameter. Likewise, if transfer of technology takes place and the technological gap narrows yet is not closed, our entire analysis still follows.

We are fully aware that factors other than technological differences impinge on migratory decisions and processes. Since the theme of technological differences and migration has been relatively neglected, our goal has been to isolate several technology-migration connections and explore their repercussions. The role of technological differences in conjunction with other factors such as information, uncertainty, and skill differentiation in the determination of labor migration is a topic of further research.

References

Bhagwati, J. N. (1964). The Pure Theory of International Trade: A Survey. *The Economic Journal, 74*, 1–84.

Bhagwati, J. N., & Srinivasan, T. N. (1983). On the Choice Between Capital and Labor Mobility. *Journal of International Economics, 14*, 209–221.

Diamond, P. A. (1965). National Debt in a Neoclassical Growth Model. *The American Economic Review, 55*, 1126–1150.

Findlay, R., & Grubert, H. (1959). Factor Intensities, Technological Progress and the Terms of Trade. *Oxford Economic Papers, 11*, 111–121.

Galor, O. (1986). Time Preference and International Labor Migration. *Journal of Economic Theory, 38*, 1–20.

Galor, O., & Ryder, H. E. (1989). Existence, Uniqueness, and Stability of Equilibrium in an Overlapping-Generations Model with Productive Capital. *Journal of Economic Theory, 49*, 360–375.

Kemp, M. C., & Kondo, I. (1986). *An Analysis of International Migration: The Unilateral Case*. University of New South Wales.

Ruffin, R. S. (1984). International Factor Movements. In R. W. Jones & P. B. Kenen (Eds.), *Handbook of International Economics* (pp. 237–288).

Wong, K. (1983). On Choosing Among Trade in Goods and International Capital and Labor Mobility: A Theoretical Analysis. *Journal of International Economics, 14*, 223–250.

Are Workers in the Developing World Winners or Losers in the Current Era of Globalization?

Nita Rudra

1 Introduction

It is crucial to understand the link between globalization and expanding opportunities for labor in less developed countries (LDCs). One group of scholars and policymakers, or the globalization "optimists," place great confidence in international markets and the purported worker benefits associated with it. LDC workers are considered "winners" in globalization, given increased growth, greater employment opportunities, and higher wages. To counter this faith, globalization pessimists argue that the existence of stagnant growth rates in some parts of the world, persistent

The author is grateful to Hayward Alker for valuable advice and input on this research project and James McGuire for generously providing access to his data. The SC1D editors and anonymous reviewers also provided extremely helpful feedback and comments.

N. Rudra (✉)
University of Pittsburgh, Pittsburgh, PA, USA

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_13

unemployment problems, recurring political repression, and growing surplus labor problems indicate that workers in LDCs are in fact "losers" during international market integration. Given this polemic, far too little attention has been paid to how global forces affect workers in LDCs, and even fewer studies have attempted to use cross-national data to assess whether international market pressures are actually the cause of deleterious outcomes for labor, as its critics suggest.

This chapter aims to investigate to what extent workers in LDCs, who are arguably the largest social class both domestically and internationally, are overall "winners" or "losers" in the global economy. This study is distinctive in that it looks beyond labor's economic benefits to assess if workers are also in a better "marketplace bargaining position" with globalization.[1] Most studies on globalization either focus solely on the economic gains and losses of labor in developing economies, or speculate how workers' bargaining power is affected without sufficient empirical justification.[2] I use a time-series cross-sectional unbalanced panel data set for 59 developing countries from 1972 to 1997 to investigate the effects of economic globalization on the skills of the LDC labor force (i.e., employment of high-skilled workers relative to low-skilled workers) and the size of the surplus labor pool. Based on these trends in labor market conditions, I deduce whether international market integration enhances or reduces their potential labor power (PLP), a variable used to approximate labor's marketplace bargaining power.[3]

[1] Silver (2003) identifies labor's marketplace bargaining power based on Erik O. Wright's (2000) analysis. It refers specifically to the *structural power* gamed by workers based on their location in the economic system and is a direct result of tight labor markets. Silver emphasizes three forms of marketplace bargaining power: (1) the possession of scarce skills, (2) low unemployment, and (3) "the ability of workers to pull out of the labor market entirely and survive on nonwage sources of income." She also discusses a*ssociational power*, which refers to "the various forms of power that result from the formation of collective organization of workers" and another form of structural power or *workplace bargaining power* that results "from the strategic location of a particular group of workers within a key industrial sector." Associational and structural bargaining powers are often positively correlated *in the long run* (e.g., teachers, European automobile workers).

[2] The edited volume by Candland and Sil (2001) is a notable exception. The case studies are limited to the experiences of five developing countries: Mexico, Brazil, India, Pakistan, and China. Silver (2003) also discusses the effects of globalization on labor, but focuses on identifying waves of labor unrest and uses the world system rather than nation states as the level of analysis.

[3] PLP is an index I developed to make inferences about the extent of labor's bargaining power. I place emphasis on *potential* labor power since, as Silver (2003) also acknowledges, workers' bargaining power does not automatically mean that workers will apply that power to demand greater political and economic benefits.

Findings in this study reveal that there are merits to the claims of both pessimists and optimists. The effects of globalization on workers in LDCs are conditional upon a nation's level of economic development. Consistent with the views of the globalization optimists, market integration has brought greater job opportunities for workers in the low-, middle-, and high-income LDCs.[4] However, as the globalization pessimists have charged, the results also indicate that globalization has exacerbated surplus labor problems in all but the high-income LDCs. Labor solidarity in the lowest-income countries has particularly come under pressure with globalization. Large populations of low-skilled workers, faced with intense competition from surplus labor, make it extremely difficult for them to overcome collective action problems in the marketplace and subsequently form broader labor alliances. The value placed on such alliances is based on Esping-Andersen's (1990) observation that working-class and white-collar coalitions have historically been most the most decisive force influencing social policies.

Based on these trends, I conclude that although workers in the lower-income LDCs are experiencing greater economic gains (greater employment opportunities), they are not necessarily increasing their bargaining power with employers and, even less likely, with the government. The repercussions may be significant: labor-friendly policies (e.g., higher wages, national welfare programs, employment benefits, political freedoms, etc.) will be inconceivable in poor nations undergoing globalization.

In the higher-income LDCs (e.g., Korea and Singapore), my findings reveal that market integration has helped improve the composition of workers' skills and has encouraged a slight reduction in surplus labor. Under these labor market conditions, I argue that workers in high-income LDCs face better prospects for mobilization. Hence, labor's economic and potential labor power in the richer LDCs is improving with globalization.

I structure the chapter by first evaluating the shortcomings of the outstanding literature on globalization and labor in LDCs. Second, I present an argument for why the effects of labor on globalization are likely to be

[4] Categories of low-, lower-middle, upper-middle, and high-income LDC are based on *World Development Indicators* 2002. Even though Greece and Mexico are also OECD countries, I included them in my LDC dataset because neither country was categorized as "high-income" for most of the time period of this study. Korea is included because it was not an OECD country until the 1990s.

conditional upon the level of economic development. Third, I use panel regression econometric techniques to assess if and how globalization affects labor in low-, middle-, and high-income LDCs. Finally, I conclude by discussing the implications of the econometric results for LDCs.

2 Globalization and Labor: The Pessimists Versus the Optimists

The debates on globalization and its effects on labor as a social actor are highly polarized. Yet little has been done to systematically assess which position the evidence favors in LDCs. On the one hand, globalization optimists apply international trade theory to argue that low-skilled labor in LDC gains with openness in terms of standard economic benefits (e.g., employment and wages). The empirical evidence supporting this position is mixed, and their models exclude any consideration of how globalization might affect labor's bargaining power in the marketplace. Conversely, pessimists who argue that international market integration is harmful to labor tend to rely on anecdotal, case-specific, or region-specific evidence, and they place primary emphasis on issues related to the weakening of labor's bargaining power (e.g., repression and poor working conditions). I argue that to achieve a sense of whether workers as a social class are improving their lot with globalization, we must assess changes in both their economic strength *and* bargaining power. A large- *N*, cross-regional assessment of the effects of international market integration can determine with more confidence whether we can generalize about particular labor patterns across LDCs.

The optimists draw from the Hecksher-Ohlin (H-O) theory of international trade, which focuses on the importance of factor endowments in determining who wins and who loses with globalization (Learner, 1996; Krugman and Venables, 1995; World Bank, 2000, 2001).[5] Because trade increases the demand for abundant factors, the theory asserts that low-skilled labor in LDCs should experience higher wages and employment with increased exposure to international markets, whereas in the more developed countries, high-skilled labor and capital should gain. This logic could also be applied to capital flows. Both productive and financial capital flows will increase in nations that are more efficiently utilizing their most

[5] This literature emphasizes that wages and income will increase with openness as long as the LDCs are fairly rapid globalizers, have strong institutions, and good governance.

abundant factor. Thus, the optimists contend that globalization improves the economic conditions of workers in labor-abundant LDCs at the expense of low-skilled labor in the capital-rich developed world. This neoclassical trade theory has become the basis of conventional wisdom on the relationship between globalization and labor.

Although such analyses help establish a conceptual framework by which to evaluate whether there is a systematic relationship between globalization and labor in LDCs, they suffer from two related shortcomings. First, the empirical evidence provided by these models is inconclusive. While several studies find that income and employment improve with globalization in LDCs (Ghose, 2000; Brunner, 2003), others determine that globalization actually worsens workers' economic position. The latter studies arrive at their conclusion by additionally considering the possible countervailing effects of structural conditions, that is, high surplus labor, lack of access to new technology, and rising global demands for skilled labor rather than low-skilled labor. Wood (1997), for example, finds very different results when he factors in characteristics of the contemporary world market such as skill-biased technological progress and the greater competitive pressures caused by the entrance of large low-income countries (e.g., China and India). His findings challenge the conventional wisdom of the optimists by demonstrating that openness has reduced the wages of unskilled labor relative to high-skilled labor in Latin America. Others such as Berman and Machin (2000), Robbins and Gindling (1997), Mazumdar and Quispe-Magnoli (2002) support Wood's contention that skill-biased technological change disadvantages low-skilled workers in LDCs. Findings from these studies emphasize the importance of taking additional domestic-level international and domestic variables into consideration.

Second, by focusing primarily on economic consequences, the aforementioned studies are arguably narrow in scope. Globalization optimists tend to focus on employment or wages as their primary dependent variable. They exclude any consideration of the bargaining prospects of labor during globalization. Even if we accept the basic premise of the H-O theory of international trade, it is yet unknown whether labor's increased economic power (i.e., higher wages and/or employment) translates into stronger bargaining power.

In contrast, the globalization pessimists downplay economic gains and focus primarily on how international markets have consequences that dampen workers' bargaining position. Often using illustrations from the

East Asian newly industrialized countries (NICs), these studies argue that globalization tends to encourage governments to be repressive toward labor to keep labor costs down and improve competitiveness in international markets. Labor's bargaining position erodes with globalization as wages, working conditions, and opportunities for mobilization and protest are forcibly suppressed (see, for example, Munck, 2002; Prasad, 1998; Manning, 1998; Beeson and Hadiz, 1998; Frenkel and Peetz, 1998; Deyo, 1989). Deyo (1989), for instance, presents the extreme subordination and exclusion of workers as the "dark side" of the Asian miracle. Such polemical accounts give rise to the growing alarm by many observers that capital is rapidly gaining rights over labor in a globalizing world.

Note that a small body of literature challenges these claims and argues contrarily that labor's bargaining power has been resilient in the face of globalization. Yet the few studies that provide empirical evidence that unions and labor mobilization strategies have *not* been adversely affected by globalization are based mostly on the experience of the Organization for Economic Cooperation and Development (OECD) nations (Scruggs and Lange, 2002; Golden et al., 1999). Frundt (2002), Webster and Lipsig-Mumme (2002), and Frenkel (1993) are exceptions that do focus on poorer nations and postulate that globalization can present greater opportunities for labor, such as incentives for mobilization and solidarity with other social groups.[6] Yet these studies lack either empirical evidence or generalizability of their thesis beyond their selected intraregional case studies. Finally, Silver focuses on the rise and decline of the world labor movement under different global political conditions. This approach, while informative, leaves out how country-specific factors might shape labor outcomes.[7]

To summarize, existing analyses have not yet considered how globalization has affected the overall circumstances of labor, and they have not presented empirical evidence that is sufficiently able to discriminate between the optimist and pessimist hypotheses. We do not have a comprehensive picture of how globalization affects labor groups across LDCs.

[6] Silver (2003) also analyzes labor and globalization, but with a different focus. Her approach centers on the interaction between global patterns of labor unrest and world political dynamics. Furthermore, globalization is considered mainly in terms of the increasing mobility of capital.

[7] Silver(2003) makes a good case for why Galton's problem precludes the necessity of observing cross-national variations. Nonetheless, this argument is not supported with empirical evidence.

How do we assess whether workers are winners in globalization? Do labor's economic gains automatically beget a stronger bargaining position under conditions of globalization? Is international market integration presenting the same opportunities to workers in relatively richer and poorer LDCs? Before a rigorous test of the pessimist and optimist hypotheses can be conducted, it is essential to first parse out the reasons why globalization might have a varying effect on workers in different type LDCs.

3 Globalization, Level of Economic Development, and Potential Gains (or Losses) for Labor

The fundamental question in this analysis is whether exposure to the global economy creates conditions in which labor is better-off economically *and* in a stronger position to demand better living and working conditions. For the purposes of this investigation, globalization refers to the expansion of international markets along three commonly recognized dimensions: trade, foreign direct investment (FDI), and portfolio flows. There are several reasons why globalization is thought to generally improve prospects for workers. A more rigorous application of existing international economic and political theories reveals that the impact of globalization on workers ultimately differs according to a nation's level of economic development.

To begin with, we should consider whether gains in workers' bargaining power are necessarily linked to the economic gains from globalization. Trade, or specifically exports of manufactured goods, is expected to encourage greater employment opportunities and eventually higher wages in LDCs, *provided that the exported goods are intensive in their most abundant factor, which is low-skilled labor*. Yet if these economic gains are realized, does this automatically translate into an improvement in labor's bargaining position, giving labor a greater "voice" in the marketplace and possibly even in policy debates (e.g., personal freedoms, welfare programs, higher wages, etc.)?[8] Conventional wisdom accepts this to be the case. As Rogowski (1989) argues, owners of locally abundant factors do in fact

[8] Silver (2003) suggests that improvements (recessions) in marketplace bargaining power do not always translate to greater (weaker) "associational power" in the short run (see fn.l). for example, British, Chinese, and Indian textile workers in the nineteenth century. However, she concludes that "these were the exceptional cases; most commonly, associational power was not sufficiently strong to compensate for the weak structural power of workers."

experience greater political power with increased trade. His study would predict that low-skilled labor in LDCs will be in a better bargaining position because of the greater wealth that accompanies openness.[9]

Yet there are several reasons to suspect that labor's bargaining position may not be enhanced by globalization, despite their economic gains. First, LDCs continue to maintain large reserves of surplus labor that do not bode well for the organizing capacity of labor. As Mancur Olson (1971) argues, tight labor markets are a necessary condition for strong bargaining power. He claims "it appears that whenever tight labor markets … increased labor's bargaining power, unions demanded and obtained union recognition and some form of compulsory membership (Olson, 1971, p. 82). Union membership has then accordingly also increased." In LDCs, labor markets are far from tight. Much of low-skilled labor is still employed in traditional nontradable (informal) sectors, which is exceptionally difficult to mobilize (ILO, 1997, 1999, World Bank, 1995).[10] In addition, it has been argued that globalization exacerbates labor market dualism. Studies have shown the "race to the bottom" for wages has been intensified by the recent entrance of large low-wage economies into global markets (Wood, 1997). Firms now have stronger incentives to reach out to the surplus labor population and take advantage of both their lower labor costs and greater labor market flexibility (Portes, 1990; Deshpande and Deshpande, 1998; Harriss-White, 1999; Papola, 1994).

The second reason that globalization may be unfavorable to labor's bargaining position is because growing numbers of low-skilled workers in developing countries can dampen labor solidarity. Low-skilled workers are initially difficult to mobilize because they have little education, work erratic hours, and a growing percentage of them are women, who according to the literature, are particularly hard to organize (Lok, 1993; Ingerson, 1984). In addition, this group is large in LDCs and growing larger with globalization. This situation results in collective-action problems because, according to Olson (1971), the larger the group, the greater the propensity for the free-rider problem to occur and the less likely the collective good (e.g., unionization) will be supplied.

[9] Conversely, capitalists and skilled workers will be better able to influence government policies in more developed countries.

[10] Nontradable informal activities include services such as haircutting, domestic help, street vending, etc.

Little systematic evidence refutes the claim that low-skilled labor populations generally have weak bargaining power. India is a paradigmatic example of how a large percentage of low-skill and surplus workers result in lower potential labor power. Low bargaining power of labor in India is particularly telling since the generous political freedom afforded to labor since independence in 1947 is contrary to the experience of most LDCs. Before globalizing, although the actual numbers belonging to unions were high, the percentage of the workforce unionized in India was a low 6.5 percent. Even then, the Indian labor movement has been marked by intense interunion rivalry that has consistently imposed limits on their collective bargaining power. After India began globalizing in the early 1990s, interviews conducted for this study with both workers and leaders of some of India's strongest unions confirmed that growing surplus labor problems were making it even harder for them to overcome their collective action problems.[11]

In contrast, *skilled* labor groups in LDCs are more capable of surmounting collective action problems in LDCs. These LDC labor groups are generally smaller in size, in low supply and high demand, and, significantly, are less threatened by a surplus labor population. In fact, the recruiting grounds for labor organizations in most LDCs have historically occurred in the skilled industries (e.g., heavy industries, white-collar workers) (Manning, 1998; Deyo, 1989). South Korea represents an important example of how a growing skilled labor population can affect labor strength. Even though labor faced a long history of repression in Korea, the labor movement flourished in the early 1980s alongside the success of the heavy and chemical industries that promoted a skilled workforce. Ultimately, the higher the numbers of skilled laborers relative to low-skilled laborers, and the lower the surplus in a country, the greater the chance that labor will be stronger in the marketplace and able to forge the type of broader coalitions discussed by Esping-Andersen (1990).

Combining H-O and Olson's theory of collective action suggests that labor in the lower-income LDCs will be the most vulnerable to employer and government hostilities as openness occurs. It is often overlooked that all LDCs do not have the same factor endowments. According to Wood (1997, p. 40), "analysts devote insufficient attention to variations among developing countries in the skill intensity of exports, which theory implies should be higher in better educated countries." Based on the logic

[11] Surplus labor is called "casual labor" in India.

outlined above, the low-income LDCs, which are primarily exporters of labor-intensive goods, are likely to experience growing numbers of low-skilled workers, persistently large pools of surplus labor, and declining bargaining power *under conditions of globalization*.[12] The opposite result is expected in the higher-income LDCs since their most abundant factor is high-skilled labor. These LDCs not only have a larger percentage of workers formally employed in the manufacturing sectors, but they also have a more skilled labor force. Literacy rates and school completion rates tend to be higher in the richer LDCs, suggesting that better bargaining arrangements for labor are more likely. In addition, many middle- and high-income countries have histories of state corporatism and continue to maintain political parties with ties to unions. Paradoxically, it is possible that governments in these countries may have indirectly aided the growth of labor movements by setting up confederations that were originally meant to control and preempt radical labor movements.[13] Given these significant differences between richer and poorer LDCs, it is a serious mistake to consider labor in developing countries as a monolithic group confronted by the challenges of globalization.

4 The Evidence

Trends in Labor Market Conditions

To obtain an initial sense of the ways in which labor in LDCs is being affected by globalization, I present trends in employment, surplus labor, and potential labor power (PLP), which is a more direct measure for labor's marketplace bargaining power and may also indicate labor's overall political power (see Appendix 2 for details). Trends in employment and wages, also constructive for evaluating H-O, are not included here because cross-country annual time-series data on these two variables are sparse for developing countries.

Scholars generally agree that although some significant economic reforms were undertaken by LDCs in the 1970s, most developing countries began opening their trade and capital markets in the early to mid-1990s (World Bank, 2000; Yusuf, 1999; Montiel, 1994; Dean, 1995). We would then expect that the impacts of globalization would be

[12] See, ILO (2002) for more details on growing levels of surplus labor.
[13] See, for example, Posusney's (1997) analysis on labor in Egypt.

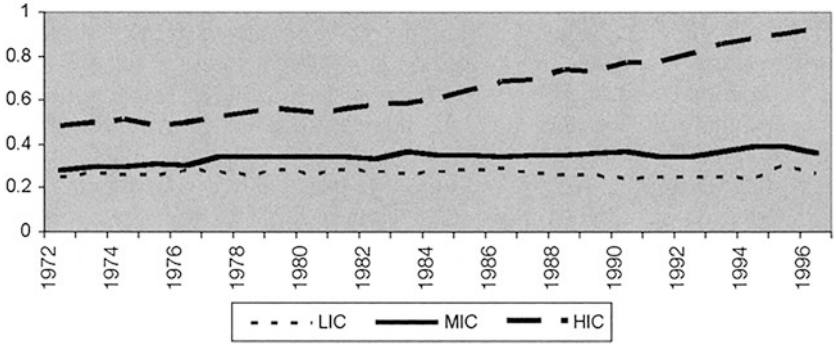

Fig. 1 Ratio of skilled labor employment/low-skill labor employment. (Source: Industrial Statistics Database, 1963–1998 (Vienna, Austria: United Nations Industrial Development Organization, 2000); Notes: See Appendix 1 for a detailed explanation of how skill ratios are calculated. "Low," "mid," and "high" refer to low-income, middle-income, and high-income LDCs as classified by the World Development Indicators, CD ROM [Washington, DC: World Bank, 2002])

most prominent after the mid-1980s. Figure 1 presents comparative trends in the ratio of high-skill employment relative to low-skill employment in low-income countries (LIC), middle-income countries (MIC), and high-income countries (HIC). Figure 1 reveals that the trend is slightly declining in the LICs, suggesting that over time, the employment in skilled labor is less than the employment in low-skilled labor. The trend for the MICs reveals that the skill/low-skill ratio begins to gradually increase only after the mid-1980s. Finally, trends in the HICs, which are generally more abundant in high-skilled labor, exhibit increasing levels of skilled-labor employment. It is interesting that the "gap" between the richest group of developing nations and the MICs and LICs began to widen *after* globalization pressures hit most LDCs in the mid-1980s.

Significantly, these trends are consistent with H-O model predictions and seem to bode well for labor in all LDCs. Employment appears to be increasing in the sectors associated with their most abundant factors, underscoring the structural differences between the three sets of countries. In the LICs, nations with the comparatively large low-skilled labor populations, the slightly downward trend in the ratio of skilled to

low-skilled labor suggests that nations with very low per capita incomes are using their factor endowments efficiently by exporting labor-intensive products commensurate with international demand. Employment of low-skilled workers increases relative to high-skilled workers. In comparison, the average growth of the ratio of high-skill to low-skill labor in MICs is 30 percent greater than it is in the LICs. The contrast with the HICs is even more pronounced with the ratio of skilled to low-skilled labor *doubling* during the time period of this analysis.

Figure 2 presents another important labor market trend. This figure shows that surplus labor has slightly declined over time. The concept of surplus labor is drawn from the development of economics literature, which claims that large surplus labor populations exist in the Middle East, parts of Asia, Latin America, Eastern Europe, and Africa.[14] The concept refers to how much "hidden" unemployment is prevalent in the economy and suggests that the supply of labor to industry is unlimited. Along with others, Amartya Sen (1966) identifies this type of labor as "hidden" in the sense that it can be removed without reducing the total amount of output produced, assuming the remaining workers would work harder (see also Wellisz, 1968). Such individuals are considered to have low (or zero) marginal productivity and are not counted as part of the official unemployed.[15]

While the declining trends in surplus labor in Fig. 2 are encouraging, data in this graph should be interpreted with caution. Surplus labor will automatically decline with the slowing of population growth, which, according to a recent United Nations report, has been unexpectedly taking place in many developing countries.[16] It is then possible to treat the skill ratio in Fig. 1 and surplus labor as independent categories. In other words, a shrinking (expanding) surplus labor pool *may or may not be* related to increasing (decreasing) employment of skilled and low-skilled labor.

In addition to the population growth rate, the surplus labor pool may vary independently of the skill ratio for two alternative reasons. The first is

[14] For more details, see Baer and Herve (1966).

[15] The logic holds that in overpopulated countries more people are employed than needed to produce output. As long as hidden unemployment prevails, its supply to industry is unlimited in the sense that industries can expand or be created without affecting the prevailing wage (See Wellisz, 1968).

[16] The United Nations Economic and Social Council Commission reports that fertility has declined in many nations since the 1970s, particularly in middle-income countries.

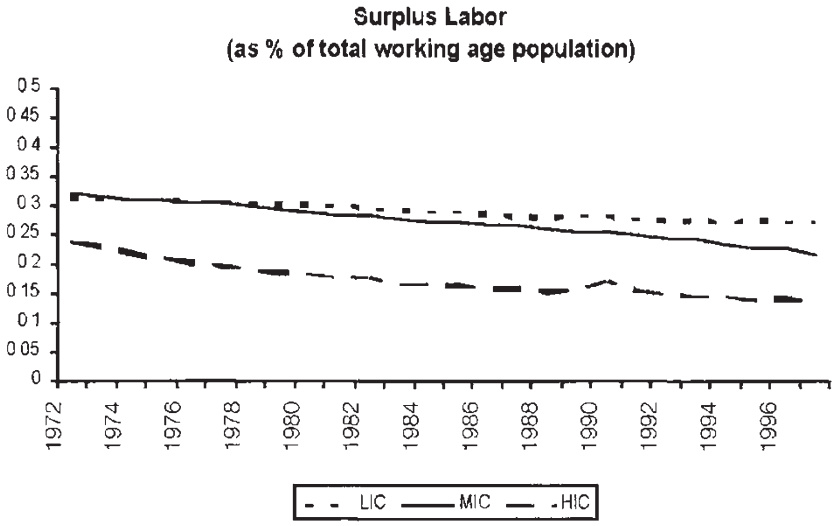

Fig. 2 Surplus labor (as % of total working age population). (Source: Industrial Statistics Database, 1963–1998) [Vienna, Austria' United Nations Industrial Development Organization, 2000]; Notes: Surplus labor is calculated as the ([working age population minus students enrolled in secondary education minus students enrolled in "post-secondary" education] minus [labor force] (/(working age population))

that reduced student enrollments in secondary and tertiary education can increase the level of surplus labor. Wood and Ridao-Cano's (1999) analysis offers insights into how and why this might occur during globalization. They demonstrate that enrollment in higher education decreases when countries begin to specialize in goods of low-skill intensity and individuals thereby see less economic advantage in advancing their education. It is feasible that the skill ratio can decrease (employing more low-skilled labor relative to high-skilled) while surplus labor is increasing. The second reason is that surplus labor will not change if the "officially" unemployed are the only ones finding the jobs.[17] In other words, if those who are already part of the labor force are filling the low-skilled or skilled positions, the skill ratio will change, but surplus labor will not.

[17] The labor force, which includes the employed and unemployed, is part of the formula for surplus labor.

Figure 2 further reveals that the average *level* of surplus is still high (i.e., 22 percent by the late 1990s, compared to 14 percent in the HICs) in both low- and middle-income LDCs.[18] It is also interesting that the level of surplus labor in the middle-income countries is only slightly lower than the low-income LDCs. This is not surprising since many researchers argue that the persistence of high amounts of surplus labor continues to be a serious problem throughout the developing world (Loayza, 1997; Kuchta-Helbling, 2000; Bangasser, 2000; Gallin, 2001; Stone et al., 1996; UN, 1992/1993). Noteworthy is that the difference between surplus labor in MICs and LICs begins to widen after globalization forces hit the LDCs in the mid-1980s, which suggests that the MICs are marginally better at absorbing excess labor under globalization conditions.

Figures 1 and 2 provide some indication that the economic conditions of labor in LDCs are improving. But what do they say about labor's bargaining situation? How do we assess if labor's ability to organize and demand better living and working conditions is similarly expanding? Measuring bargaining power is challenging in LDCs. The most common method of assessing labor power is by unionization rates (see, for example, Deverajan et al., 1997; Western, 1997; Galenson, 1962, 1994). A fundamental problem with this indicator is that, unlike the developed countries, union density (percent of the working population that is unionized) is not comparable across LDCs. Many LDC governments mandate compulsory membership in corporatist unions and impose constraints on labor's demands, leadership, and internal governance. China, for example, has the highest union density in the developing world, but labor has very little bargaining power (Frenkel and Peetz, 1998; Chan and Senser, 1997). In general, unionization rates exaggerate labor's independent political strength in LDCs.[19]

Given the unreliability and weakness of direct organizational measures, I establish PLP that gives a sense of labor's marketplace bargaining power by combining two direct measures of structural conditions in labor markets. The index is motivated by the observation that the bargaining potential of labor is likely to increase with the ratio of skilled to unskilled workers,

[18] This can be compared to an average of 7 percent of surplus labor in OECD countries and 13 percent in the HICs.

[19] See for example, Valenzuela, 1989; Banuri & Amadeo, 1991. McGuire, 1997 adds that unreliability of union data can result in huge discrepancies in existing cross-country compilations of union-density estimates.

given the greater labor's capacity for collective action,[20] and decreasing with the size of "surplus" labor. PLP is based on Silver's (2003) *marketplace bargaining power*, which is the power that "results directly from tight labor markets" and Olson's (1971) *Logic of Collective Action*. Based on Olson, we know that size and incentive matter for collective action and, building on Silver, labor's marketplace bargaining power in LDCs is influenced by the proportion of skilled and low-skilled workers and surplus labor.

If labor is in a better position *vis-à-vis* the market, can we assume that they have more bargaining power in policy debates? Appendix 2 illustrates in detail how and why PLP can also be used to make inferences about labor's political power. Most important, PLP shows a fairly strong correlation (0.61) with James McGuire's (1999) labor strength index (LSI).[21] To provide a *direct* measure of political power, PLP would need to include other measures such as unions' connections to parties or government. This is an effort reserved for a future study on labor's political bargaining power.

To create an index, each country's score was divided by the highest value in my larger data sample (i.e., Sweden = 87) and multiplied by 100. Assuming that there is always some surplus labor and some low-skilled laborers, the PLP measure is:

$$\mathrm{PLP} = \left[\frac{\left(\text{Number of skilled workers} / \text{Number of low-skilled workers}\right) \times \left(1 / \text{Surplus labor as\% of working age population}\right)}{87} \right] \times 100$$

PLP falls as the ratio of low-skilled workers to skilled workers increases, *and* as surplus labor rises. As surplus labor shrinks and labor markets become tighter, PLP increasingly depends on the ratio of skilled to low-skilled workers. To summarize, *all other things being equal*, labor will be in a more favorable bargaining position under the following scenarios: (1)

[20] For more detailed hypotheses on why low-skilled labor groups in LDCs are difficult to organize, see Deyo (1989); Gereffi (1995); Ingerson (1984).

LSI covers mostly mid-1990s data.

[21] The decade dummies were dropped from the model if they were insignificant across models.

the percentage increase in high-skill workers is greater than the percentage increase in surplus labor (surplus labor may also be decreasing or unchanging); (2) surplus labor decreases and the skill ratio remains relatively constant (labor markets are tighter); and (3) the percentage increase in low-skilled workers is accompanied by the equal or greater percentage decrease in surplus labor.

This assessment of PLP is limited to the manufacturing sector, since data are not available for most countries outside this sector. Significantly, although there have been notable exceptions (e.g., banana workers in Honduras), research has shown that most labor-organizing activities began to occur in this sector. This pattern is common to both developed and developing countries.

Tables 1, 2, and 3 reveal how countries rank on PLP, yet the measure of PLP should be viewed with some caution. It does not capture institutional and other political factors that may mediate between the economic variables and labor's capacity for collective action. A country may have high PLP, for example, but mobilization may be discouraged by a

Table 1 Potential labor power in low-income countries

	PLP	*PLP*	*PLP*	Δ*PLP*
Country	1970s	1980s	1990s	1970s–1990s
Kenya	2.37	1.95	1.72	-0.65
India	2.24	2.77	3.65	1.41
Ghana	1.72	2.05	2.13	0.41
Zambia	1.33	1.62	1.31	-0.02
Nicaragua	1.24	1.11	1.30	0.06
Cameroon	0.98	0.61	0.34	-0.64
Indonesia	0.90	1.19	1.26	0.36
Zimbabwe	0.89	1.11	1.05	0.15
Tanzania	0.82	0.75		
Pakistan	0.82	0.97	0.52	-0.30
Mozambique	0.77	1.11	1.32	0.54
Bangladesh	0.58	0.68	0.43	-0.15
Malawi	0.57	0.81	0.94	0.37
Nigeria	0.56	0.98	1.02	0.46
Liberia	0.42	1.83		
Mali	0.38	0.40	1.19	0.81
Nepal	0.28	0.16	0.85	0.57
Lesotho		0.28	0.11	
Average	0.99	1.13	1.20	0.23

Table 2 Potential labor power in middle-income countries

	PLP	*PLP*	*PLP*	Δ*PLP*
Country	1970s	1980s	1990s	1970s–1990s
China	11.05	11.56	12.08	1.03
Mexico	2.60	4.47	5.72	3.12
Argentina	2.33	2.53	3.12	0.79
Turkey	2.23	1.91	2.10	-0.12
Malaysia	2.03	3.20	4.84	2.81
South Africa	2.03	2.62	3.20	1.17
Brazil	1.69	1 92	2.80	1.12
Trinidad and Tobago	1.63	1.90	1.96	0.33
Paraguay	1.61	1.55	1.48	-0.13
Philippines	1.56	1.73	4.02	2.46
Chile	1.45	1.19	1.43	-0.03
Venezuela	1.39	1.47	1.55	0.15
Thailand	1.39	2.14	6.57	5.18
Sri Lanka	1.19	0 46	0.53	-0.66
Jordan	1.09	0.95	0 94	-0.15
Colombia	1.05	1 46	2.22	1.17
Uruguay	1.02	1.47	2.54	1.52
Iran	0.94	1.25	1.80	0.86
Egypt	0.93	1.42	1.99	1.06
Ecuador	0.91	1.28	1.32	0.40
Mauritius	0.89	0.27	0.61	-0.28
Morocco	0.84	0.92	0.85	0.01
Tunisia	0.79		1.21	0.42
Fiji	0.78	0.69	0.41	-0.37
Panama	0.72	1.02	0.89	0.16
Dominican Republic	0.67	0.93		
Guatemala	0.66	0.95	0.99	0.33
El Salvador	0.59	0.79	0.81	0.22
Bolivia	0.58	0.75	0.96	0.37
Honduras	0.38	0.48	0.29	-0 09
Guyana	0 32			
Syrian Arab Republic	0.30	0.88	0.82	0.52
Botswana	0.08	0.33	0.63	0.55
Peru		1.57	1.39	
Costa Rica		0.94	1.17	
Average	1.45	1.73	2.22	0.77

repressive regime, state measures to "buy off" labor, such as by offering higher wages and benefits, or even a more equitable distribution of income.

Nevertheless, the advantage of using this proxy is that it can capture variations in labor's marketplace bargaining power, irrespective of the type

Table 3 Potential labor power in high-income countries

	PLP	*PLP*	*PLP*	Δ*PLP*
Country	1970s	1980s	1990s	1970s–1990s
Singapore	7.41	11.67	13.16	5.75
Israel	2.63	3.80	4.26	1.63
Korea	2.60	5.64	10.45	7.85
Greece	1.64	1.73	2.13	0.49
Kuwait	1.58	2.49	2.83	1.25
Cyprus	0.76	1.04	1.70	0.94
Average	2.77	4.39	5.75	2.98

of state-labor relations. It abstracts from the historical complexities behind the relationship between LDC governments and labor, a relationship that ranges widely between state control of labor or state corporatism (such as in Brazil, Mexico, Egypt, Ghana, and Tanzania), state repression of labor (such as in Korea, Zambia, Taiwan, and Singapore), to relative autonomy of labor (such as in India and Costa Rica). Alternative, standardized cross-country, time-series measures of labor's bargaining power for many LDCs do not exist.

Figure 3 illustrates the correlation between unionization rates and PLP in 1995. As expected, the correlation is not very strong, but overall, it is fair to conclude that most LDCs have both low unionization rates and low PLPs. Several LDCs are excluded to avoid clustering and enhance readability of the graph.

Tables 1, 2, and 3 indicate that average PLP in the high-income countries is *more than five times greater* than the average PLP in the low-income countries, and almost twice as high as the middle-income countries (in the 1990s). This is significant, particularly since the high PLP values for China drive the total average upward for the middle-income countries. It also important that 13 LICs and MICs experienced a fall in PLP over the decades as compared to only positive PLP values in the high-income countries. Once again, the low level of PLP (compared to Sweden's 87 PLP value) in the three sets of LDCs should be emphasized. Most striking, in 25 years, the average increase in PLP for the low-income countries is not even close to half of one point (.23).

In sum, a first look at the data presents a mixed picture of worker conditions in LDCs. Workers seem to gain economically in LICs and

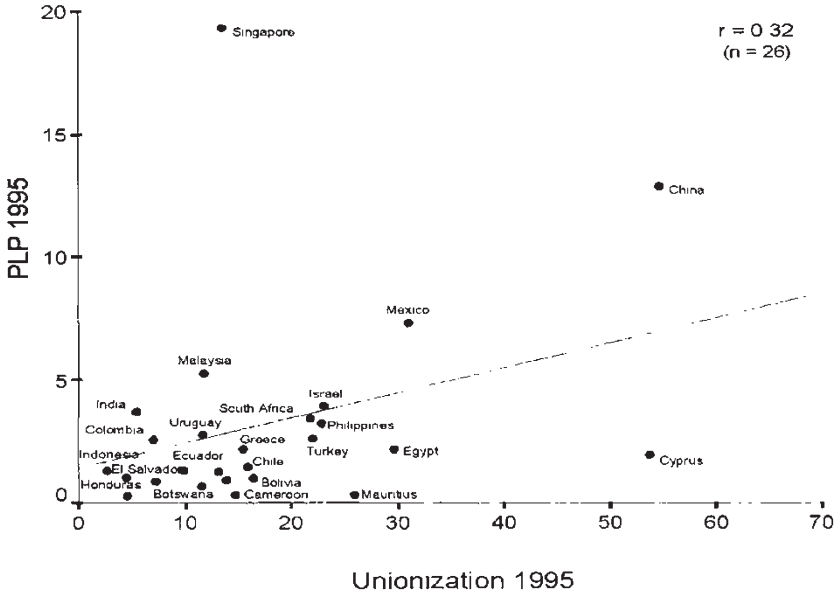

Fig. 3 Correlation between PLP and unionization (1995). (Source: International Labor Organization; Note: Unionization rates represent percent of the nonagricultural labor force that belongs to unions)

MICs, but not necessarily in bargaining terms (low PLP). However, labor in countries at relatively higher levels of development appears to be in a comparatively better economic position and marketplace bargaining power. Specifically, workers in LDCs at all income levels appear to gain significant economic benefits over time, such as greater employment, and even a minor reduction in surplus labor. Yet the absolute level of surplus in both middle- and low-income LDCs remains disturbingly high. While increases in overall PLP in MICs and LICs occur, this improvement is slight. Labor's marketplace bargaining power in the HICs seem to be the most encouraging. How much is globalization affecting these trends? Are the trends improving or worsening as a consequence of integration? Does globalization have a differing effect depending on the income category of the LDC? An econometric test will provide some answers to these questions.

5 The Model

A pooled time-series model that includes both economic and political variables affecting labor enables a more comprehensive picture of labor outcomes. I use both country dummy variables and decade dummies to control for country-specific and time-specific fixed effects.[22] Based on econometric techniques advocated by Beck and Katz (1995), I correct for both panel heteroskedasticity and spatial contemporaneous autocorrelation. In addition, problems of potential serial autocorrelation within each panel are addressed by estimating and adjusting for a panel-specific AR(1) process. This model follows Christopher Achen's (2000) recommendation against applying the standard practice of using a lagged dependent to correct for serial autocorrelation. These results provide Prais-Winsten coefficients with panel corrected standard errors (PSCE).

The lack of data limits the sample size to 59 developing countries. Nevertheless, the sample is regionally diverse and includes six high-income LDCs, 35 middle-income LDCs, and 18 low-income LDCs. The LDC labor models are as follows:

The LDC Labor Models

$$\begin{aligned}
&\text{Skill}_{it} = b_1\text{Xmanuf}_{it-1} + b_{2:}\text{port}_{it-1} + b_3\text{fdi}_{it-1} \\
&+b_4\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{LIC}_{\mathbf{it}} + b_5\mathbf{port}_{it-1}{}^{*}\mathbf{LIC}_{\mathbf{it}} \\
&+b_6\mathbf{fdi}_{it-1}{}^{*} + \mathbf{LIC}_{\mathbf{it}} + b_7\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{MIC}_{\mathbf{it}} \\
&+b_8\mathbf{port}_{it-1}{}^{*}\mathbf{MIC}_{\mathbf{it}} + b_9\mathbf{fdi}_{it-1}{}^{*}\mathbf{MIC}_{\mathbf{it}} \\
&+b_{10}LIC_{it} + b_{11}MIC_{it} + \sum\left(b_j X_{jit}\right) \\
&+\sum\left(b_k \text{decade}_{kt}\right) + \mu_{it}
\end{aligned} \tag{1}$$

[22] Estimating a measure of first-order correlation (p) has the effect of making the time-series data stationary. This measure is used to transform all variables in the model according to the formula: $y^*_t = y_t - (\rho^* y_{t-1})$. This partial differencing reduces suspicions about spurious results, particularly the concern that the effects of potentially high degree of institutional inertia exhibited by the dependent variable are not captured.

$$\begin{aligned}\text{Surplus}_{it} &= b_1\text{Xmanuf}_{it-1} + b_{2:}\text{port}_{it-1} + b_3\text{fdi}_{it-1} + b_4\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{LIC_{it}} \\ &+ b_5\mathbf{port}_{it-1}{}^{*}\mathbf{LIC_{it}} + b_6\mathbf{fdi}_{it-1}{}^{*} + \mathbf{LIC_{it}} + b_7\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{MIC_{it}} \\ &+ b_8\mathbf{port}_{it-1}{}^{*}\mathbf{MIC_{it}} + b_9\mathbf{fdi}_{it-1}{}^{*}\mathbf{MIC_{it}} + b_{10}LIC_{it} + b_{11}MIC_{it} \\ &+ \sum\left(b_j X_{jit}\right) + \sum\left(b_k \text{decade}_{kt}\right) + \mu_{it}\end{aligned} \tag{2}$$

$$\begin{aligned}\text{PLP}_{it} &= b_1\text{Xmanuf}_{it-1} + b_{2:}\text{port}_{it-1} + b_3\text{fdi}_{it-1} + b_4\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{LIC_{it}} \\ &+ b_5\mathbf{port}_{it-1}{}^{*}\mathbf{LIC_{it}} + b_6\mathbf{fdi}_{it-1}{}^{*} + \mathbf{LIC_{it}} + b_7\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{MIC_{it}} \\ &+ b_8\mathbf{port}_{it-1}{}^{*}\mathbf{MIC_{it}} + b_9\mathbf{fdi}_{it-1}{}^{*}\mathbf{MIC_{it}} + b_{10}LIC_{it} + b_{11}MIC_{it} \\ &+ \sum\left(b_j X_{jit}\right) + \sum\left(b_k \text{decade}_{kt}\right)\sum + \mu_{it}\end{aligned} \tag{3}$$

Skilled-labor employment (ratio of skilled/low-skilled), surplus labor (as % of working age population), and PLP are the dependent variables that attempt to capture labor's economic and bargaining power, respectively. The *bs* are parameter estimates; *LIC*, *MIC*, and *HIC* are dummy variables that represent the income level of the countries; m is an error term; ΣX represents the vector of control variables. The globalization variables, that is, exports of manufactured goods (Xmanuf), portfolio flows (port), and FDI are lagged to account for the period of adjustments. Each of the globalization variables is interacted with the income dummies to disentangle the effects of international markets on labor in countries at different levels of economic development (*LICs, MICs, HICs*). See Appendix 1 for detailed explanations of the source and measurement of variables.

Including both FDI and exports of manufactured goods might cause endogeneity problems since researchers argue that FDI is likely to lead to greater exports and vice-versa. The conventional method for coping with this endogeneity issue and mitigating bias of the regression estimates is to use the instrumental variable approach. The difficulty is to find outside data (or instruments) that are uncorrelated with the error of the equation and, at the same time, highly correlated with the explanatory variables. Therefore, instead of using conventional variables as instruments, the Lewbel (1997) procedure of using higher moments of the FDI variable as the instrument is applied in this model.

6 Variables

The Dependent Variables: PLP, Ratiosk, and Surplus Labor

Assessing the effects of international market indicators on three labor variables presents a more comprehensive picture of how globalization affects workers in LDCs. Each component captures different labor effects. The ratiosk variable and the surplus labor variables indicate whether workers' economic conditions improve in response to globalization. The PLP variable is the rough proxy of labor's marketplace bargaining power and suggests the extent of labor's organizing potential.

Globalization: Xmanuf, Portfolio, FDI

The primary objective of this analysis is to test if and how deepening market integration affects labor in LDCs. To incorporate the primary international economic forces said to affect workers, I measure the degree of globalization by the exports of manufactured goods and portfolio and FDI flows (as percentage of GDP). I do not apply the conventional method for assessing trade openness by measuring exports plus imports relative to GDP for two reasons. First, exports of *manufactured* goods are a more appropriate test of H-0 in the current era, particularly since it is widely expected that the "export boom" in labor-abundant LDCs will occur in this sector (see Lall, 1998). Second, developing countries have increasingly focused on expanding their manufactured exports because compared to primary products, it is believed that the former is associated with greater spillover effects and contributes to superior growth performance (e.g., employment opportunities, spread of technology, skills). A closer look at the data confirms that Xmanuf is the fastest growing and largest component of the trade variable in developing countries. Emphasizing manufactured exports in this analysis may best capture LDC efforts to eliminate the anti-export bias that predated globalization.[23] The effects of capital flows in this model are disaggregated to account for the potential effect of different types of funds on labor. FDI, where foreign firms may start new production, is likely to have a more positive and direct

[23] As a robustness check, I ran the model using total exports of goods, services, and income in place of manufactured exports. The primary results were unaffected.

impact on labor than portfolio investment, which is more susceptible to herd instinct and panic.

Each of the international market variables is interacted with the LDC income variables to assess whether different levels of income matter (Xmanuf* *L/C*, Port**LIC*, *FDI***LIC* Xmanuf* *MIC*, Port**MIC*, *FDI***MIC*).[24] See Tables 4, 5, and 6 for a summary of the expected signs. Recall that the globalization optimists argue, based on the H-O model,

Table 4 Predicted signs of conditional coefficients for high-skill/low-skill ratio

Optimists				*Pessimists*			*Author*		
Income Level	Xmanuf	Port	FDI	Xmanuf	FDI	Port	Xmanuf	FDI	Port
LIC	–	–	–	0	0	0	–	–	–
MIC	- or +	- or +	- or +	0	0	0	- or +	- or +	- or +
HIC	+	+	+	0	0	0	+	+	+

Table 5 Predicted signs of conditional coefficients for surplus labor

Optimists				*Pessimists*			*Author*		
Income Level	Xmanuf	Port	FDI	Xmanuf	FDI	Port	Xmanuf	FDI	Port
LIC	–	–	–	+	+	+	+	+	+
MIC	–	–	–	+	+	+	- or +	- or +	- or +
HIC	–	–	–	+	+	+	–	–	–

Table 6 Predicted signs of conditional coefficients for PLP

Optimists				*Pessimists*			*Author*		
Income Level	Xmanuf	Port	FDI	Xmanuf	FDI	Port	Xmanuf	FDI	Port
LIC	*0*	*0*	*0*	–	–	–	–	–	–
MIC	*0*	*0*	*0*	–	–	–	- or +	- or +	- or +
HIC	*0*	*0*	*0*	–	–	–	+	+	+

Note: + represents an expected positive effect on the dependent variable; - represents an expected negative effect; 0 represents an expected insignificant or null relationship

[24] HIC interactions are excluded to avoid linear dependency.

that openness brings positive economic benefits to labor. The first column represents this optimistic view. Globalization is expected to reduce surplus labor in all LDCs and improve employment in sectors associated with their most abundant factors. If they are correct, the globalization variables will display a negative sign for skill ratios in the LICs (since low-skill populations are high and expected to increase with globalization), either a positive or negative sign in MICs (since there is a sizeable population of both types of workers, the direction of the sign is difficult to predict), and a positive sign for HICs (since there is a relatively large high-skill population). In Table 4, for example, a negative sign in the LICs confirms that international demand for low-skilled labor relative to skill labor is increasing and so H-O is confirmed (the optimists).

In the surplus labor model (Table 5), the optimists predict that globalization variables will reduce surplus labor in all countries, resulting in a negative sign regardless of income category. The expected impact of globalization on PLP is set at "0" in Table 6, which suggests no significant impact of the globalization variables on marketplace bargaining power since optimists do not hypothesize about the direct effects of international markets on labor's bargaining position.

The middle columns of Tables 4, 5, and 6 represent the pessimists' predictions. If the globalization pessimists are correct, then the conditional coefficients will reveal that globalization will reduce PLP and increase surplus labor in all LDCs. The effect of integration on the skill ratio is "0" for all LDCs because these studies tend to draw generalizations about all labor groups based on the effects of globalization on one group of workers, typically those that are part of the surplus population. In other words, if globalization increases demands for surplus labor, this tends to carry more emphasis in these analyses than the potential positive effects on formal sector employment.

Finally, the last column in Tables 4, 5, and 6 represents the effects of globalization as predicted in this analysis. If the arguments presented in this chapter are correct, in contrast with the positions of the optimists and pessimists, we can expect that the overall impact of globalization will depend on the income level of the LDC. Specifically, market expansion is likely to have beneficial *economic* effects in all LDCs, but decisively positive impacts on marketplace bargaining power in only the HICs. In LICs, with globalization, international demand for low-skilled labor is likely to encourage employment of low-skilled labor and, at the same time, either

expand or maintain the surplus labor population.[25] The latter occurs when employers have strong incentives to cut labor costs and there is a sizeable population willing to work for long hours, low wages, and minimal social protections.

At the same time, I anticipate collective action problems in LICs (represented by PLP) will be exacerbated. The effects of globalization on both the economic variables and PLP in MICs are more difficult to predict since the skill ratio is mixed and surplus labor trends show a steady decline after the mid-1980s, yet still remain at high levels (Fig. 2). There is less ambiguity about the HICs because the high-skill labor population is relatively large and surplus labor in HICs is proportionately smaller. It is easy to see why expanding job opportunities with globalization conditions could increase opportunities for higher-skilled labor in the HICs, substantially reduce the size of their surplus labor and thereby have a positive effect on PLP.

Control Variables: Democracy, Dependents, Growth, Urban, Human Capital Spending

Including control variables helps isolate the main relationship between globalization and labor while checking for other influences. PLP is expected to increase with democracy, lower dependents (suggesting that a larger percentage of the population is formally employable), higher growth, and the level of urbanization. Conversely, surplus is expected to increase with lower democracy, higher dependents, lower growth, and rapid urbanization. The controls for skill ratio are slightly different since there is no reason to expect that regime type or the number of dependents affects the skill composition of the workforce. Instead, human capital spending (per capita) is added as a control since higher government spending on education and health is apt to result in a more skilled labor force.[26]

[25] With globalization, both can simultaneously occur if incentives to work in the informal sector is greater than the incentive to enroll in school, and the increase in low-skill employment is drawn from the existing labor force.

[26] The per capita human spending is an instrumental variable (using the Lewbel procedure) in this model because of endogeneity—LDCs with large skilled labor populations are apt to invest in human capital. This variable is not included as a control for surplus labor since higher government spending on education and health in LDCs does not necessarily benefit the poor and tends to be disproportionately allocated to the middle and upper classes. See *World Development* 2000–2001.

Also, the percentage of youth (age 0 to 14) is substituted for dependents since countries with large youth populations are most likely to enjoy comparative advantage in low-skilled labor.

The Results

Equations (1), (2), and (3) were estimated for 59 LDCs, from 1972–1997, using the fixed effects procedure. The empirical findings confirm expectations in this analysis that the effects of globalization on labor are contingent upon the level of economic development. In sum, four interesting patterns emerge from the results. First, consistent with the predictions of the globalization optimists, labor's economic situation tends to improve in all LDCs, that is, the employment of workers associated with a nation's abundant factors increases with openness in countries at different levels of development. Second, the findings confirm that it is important to disaggregate the globalization variable. Exports and FDI flows have the greatest effect on workers in LICs and MICs, while portfolio flows have the strongest impact on workers in the HICs. This should not be surprising since many HICs are moving into the post-industrial stage and we can expect exports of manufactured goods to have proportionately less impact on workers.[27] The third finding, as the pessimists predict, is that international market expansion has exacerbated the surplus labor problem in the LICs and MICs. Finally, under globalizing conditions, PLP improves most decisively in the HICs. The results for PLP are less encouraging in the MICs and LICs.

Table 7 reports the regression coefficients. This model suggests that greater overall benefits should accrue to labor at higher levels of economic development. Because the globalization coefficients in Table 7 represent the relationship between market expansion and labor when MIC and LIC equal zero, it describes the situation of the high-income LDCs. The statistically significant and positive coefficients for the uninteracted portfolio and FDI variables are evidence that, as expected, marketplace bargaining power and economic conditions for workers improve with globalization in high-income LDCs.

Tables 8, 9, and 10 further exhibit tests of whether the globalization coefficients are significant at different levels of LDC income by calculating

[27] Portfolio flows in HICs are almost twice as high as flows in MICs, and almost four times the level of flows in LICs (averaged from 1972–1997).

Table 7 Fixed effects regression estimates: the relationship between globalization and labor in LDCs

Variables	*Skilled / Low skilled*	*Surplus labor*	*PLP*
Manufactures exports$_{t-1}$, * *LIC*	-0.001	-0.194	0.184
	(0.009)	(0.214)	(0.326)
FDI$_{t-1}$ **LIC*	-0.187**	5.182**	-7 903***
	(0.085)	(2.056)	(2.978)
Portfolio flows$_{t-1}$, * *LIC*	-0.032**	0.626*	-1.773***
	(0.017)	(0.354)	(0.419)
Manufactures exports$_{t-1}$ * *MIC*	-0.0005	-0.274	0.205
	(0.009)	(0.214)	(0.326)
FDI$_{t-1}$ * *MIC*	-0.177**	1.982	-7.699**
	(0.084)	(1.990)	(2.979)
Portfolio flows$_{t-1}$ * *MIC*	-0.040***	0.516^{+}	-1.756***
	(0.015)	(0.333)	(0.418)
Manufactures exports$_{t-1}$	0.003	0.238	-0.191
	(0.009)	(0 214)	(0.326)
Portfolio flows$_{t-1}$	0.039***	-0.521[a]	1 758***
	(0.015)	(0.333)	(0.418)
FDI-*$_{t-1}$	0.128	(0.089)	-2.767
(1.993)	7.716**	(2.980)	
Democracy	–	-0.068*	0.005
		(0.038)	(0.011)
Depend[b]	-0.004**	23 14***	-3 897***
	(0.002)	(2.179)	(0.588)
Urban	-0.003	0.064***	0.005
	(0.004)	(0.013)	(0.004)
Growth	-0.0003	-0.003	-0.003
	(0.0004)	(0.010)	(0.003)
Human capital spending	0.2444*	–	–
	(0.135)		
LIC	-0.027	1.154	5.046***
	(0.259)	(2 518)	(0.667)
MIC	-0.414	7.838***	3.969***
	(0.420)	(2.365)	(0.664)
HIC	-0.630	-24.63	22.70
	(0.988)	(19.35)	(29.06)
N	581	834	685
R^2	0.664	0.940	0.846

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$; $^{+}p < 0.15$

Figures in parentheses are standard errors; intercept suppressed. Decade dummies are not shown here

[a]represents instrumental variables

[b]Note that the percentage of youths was used in place of dependents for the high-skill/low-skill model

Table 8 Conditional effects of exports of manufactured goods, portfolio flows, and FDI on *skilled / low-skilled* (*Calculated from Table 3*)

LDC income category	*Exports of manufactured goods*	*FDI*	*Portfolio flows*
LIC	0.002**	-0.069*	0.008
	(0.00079)	(0.032)	(0.008)
MIC	0.003***	-0.048*	-0.0007
	(0.0009)	(0.028)	(0.0008)
HIC	0.003	0.128†	0.039***
	(0.009)	(0.088)	(0.015)

Table 9 Conditional effects of exports of manufactured goods, portfolio flows, and FDI on *surplus (Calculated from Table 3*)

LDC income category	*Exports of manufactured goods*	*FDI*	*Portfolio flows*
LIC	0.044***	2 42***	-0.105
	(0.012)	(0.004)	(0.136)
MIC	0.036***	-0.784***	-0.0004
	(0.0088)	(0.156)	(0.014)
HIC	0.238	2.77	-5.21†
	(0.214)	(0.020)	(3.33)

Table 10 Conditional effects of exports of manufactured goods, portfolio flows, and FDI on *PLP* (*Calculated from Table 3*)

LDC income category	*Exports of manufactured goods*	*FDI*	*Portfolio flows*
LIC	-0.007***	-0.187**	-0.015
	(0.002)	(0.076)	(0.038)
MIC	0.014***	0.017	0.002
	(0.004)	(0.059)	(0.004)
HIC	-0.191	7.71**	1.76***
	(0.326)	(2.98)	(0.418)

Note: Tables 8, 9, and 10: Conditional coefficients with conditional standard errors are in parentheses
***$p < 0.01$; **$p < 0.05$; *$p < 0.10$., †$p < 0.15$

the conditional coefficients for each relevant LDC combination. By substituting the appropriate values of the income variables into Eqs. (1)–(3), we can assess how international market integration affects labor in low-, middle-, and high-income countries. For example, the conditional

coefficient for the effect of globalization variables on the skill ratio in LICs can be determined by substituting 1 for LIC and 0 for MIC.

As discussed earlier, the percentage of youths was used in place of dependents for the skill/low-skilled model

$$\begin{aligned}\text{Skill}_{it} &= b_1\text{Xmanuf}_{it-1} + b_{2:}\text{port}_{it-1} + b_3\text{fdi}_{it-1} + b_4\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{LIC}^{*}(\mathbf{1}) \\ &+ b_5\mathbf{port}_{it-1}{}^{*}\mathbf{LIC}^{*}(\mathbf{1}) + b_6\mathbf{fdi}_{it-1}{}^{*}\mathbf{LIC}^{*}(\mathbf{1}) + b_7\mathbf{Xmanuf}_{it-1}{}^{*}\mathbf{MIC}^{*}(\mathbf{0}) \\ &+ b_8\mathbf{port}_{it-1}{}^{*}\mathbf{MIC}^{*}(\mathbf{0}) + b_9\mathbf{fdi}_{it-1}{}^{*}\mathbf{MIC}^{*}(\mathbf{0}) + b_{10}LIC_{it} + b_{11}MIC_{it} \\ &+ \sum\left(b_j X_{jit}\right) + \sum\left(b_k decade_{kt}\right) + \mu_{it}\end{aligned}$$

If we want to know specifically how exports of manufactured goods affect the skill ratio in LICs, the above equation simplifies to:

$$\begin{aligned}&\partial\text{skill} / \partial\text{Xmanuf} = b_1 + \mathbf{b_4}^{*}(\mathbf{1}) + \mathbf{b_7}^{*}(\mathbf{0}) \\ &\text{or simply } b_1 + \mathbf{b_4}\end{aligned}$$

The impact of exports of manufactured goods that affects the skill ratio in MICs can be determined by:

$$\begin{aligned}&\partial\text{skill} / \partial\text{Xmanuf} = b_1 + \mathbf{b_4}^{*}(\mathbf{0}) + \mathbf{b_7}^{*}(\mathbf{1}) \\ &\text{or simply } b_1 + \mathbf{b_7}\end{aligned}$$

Finally, the effect of exports on HICs is expressed as:

$$\begin{aligned}&\partial\text{skill} / \partial\text{Xmanuf} = b_1 + \mathbf{b_4}^{*}(\mathbf{0}) + \mathbf{b_7}^{*}(\mathbf{0}) \\ &\text{or simply } b_1\end{aligned}$$

Estimated for each of the globalization variables in this way, Tables 8, 9, and 10 report the conditional coefficients and their associated standard errors based on the results in Table 7. It is interesting to compare these results to the predicted signs presented in Tables 4, 5, and 6 because the findings verify greatly the predictions of this analysis. Results based on each LDC income category are explained in more detail below.

Results for LICs

Globalization brings greater economic benefits for workers in LICs through exports of manufactured goods and FDI flows, but it weakens their marketplace bargaining power (or PLP). It is not surprising that portfolio flows showed no impact on workers since this type of capital flow still represents a very small percentage of GDP in the LICs. The negative coefficient for FDI in Table 8 indicates that this type of investment employs more low-skilled workers relative to high-skill labor. At the same time, as the pessimists predicted, increasing exports and FDI fuel the expansion of the surplus labor pool and dampen PLP. Recall, the pessimists posit that surplus labor will increase because with globalization, firms have greater incentives to avoid labor market regulations and seek lower labor costs (Portes, 1990). The positive coefficient for surplus labor in the LDCs lends some support to this hypothesis. PLP is consequently weakening because the increasing numbers of low-skilled workers and larger surplus labor pools negatively affect labor solidarity.

One unexpected finding is the positive coefficient for exports in Table 8. This positive coefficient defies the predictions of the H-O model. Why would this type of market integration increase the demand for skilled labor relative to low-skilled labor in low-skill labor abundant countries? We must emphasize that the effect of exports on the skill ratio is weak (i.e., size of the coefficient 0.002) in comparison to the impact of FDI (-0.07). A possible explanation is that the inclusion of India is biasing the results for exports of manufactured goods. India engaged in one of the longest periods of economic isolationism to protect and cultivate their high-skill sectors. Although its low-skill employment has recently expanded, the overall employment of skilled workers in India is high compared to other countries at comparable levels of development. To check this possibility, I dropped India from the sample and re-ran the regressions (results not shown here). The new results reveal that exports of manufactured goods no longer have a significant effect on the skill ratio, but FDI sustains its negative and significant impact.

Breakdown of Results for MICs

The statistical findings for the MICs suggest that overall labor is slightly better off under conditions of globalization than the LICs. Similar to the LICs, the effects of globalization on workers occur mainly through exports and FDI. Exports of manufactured goods increase employment of high-skilled

labor in middle-income countries compared to the employment of low-skilled workers. This makes sense because MICs have higher literacy rates, relatively greater access to technology, and higher levels of industrialization than the LICs. Similar to India, many countries in this group (e.g., Argentina, Brazil, Venezuela, Malaysia, Thailand) engaged in long periods of industrial substitution industrialization (ISI) where governments encouraged the development and exports of their more skill-intensive sectors. In contrast, the results suggest that FDI in the MICs tends to be concentrated in labor-intensive industries that lead to growing employment of low-skilled labor.

The effects of FDI and exports on surplus labor in MICs show contrasting results. The negative and significant coefficient for FDI indicates that foreign productive investment helps absorb much of the surplus labor population. Yet interestingly, increasing exports appear to encourage both higher-skill ratios and greater surplus labor pools. Predictably, if MICs do not provide sufficient job retraining and relocation programs, low-skilled workers will be out of work as the economy moves toward higher-skilled production. Nevertheless, the size of the export coefficients suggests that the positive impact of exports on surplus labor is small, particularly when compared to the negative effects of FDI on surplus. This finding is consistent with Fig. 2, which shows that surplus labor declined more rapidly in these countries after policies supporting globalization were adopted. Regardless, these contrasting findings raise interesting questions and deserve further future investigation.

Finally, the PLP findings in Table 10 are mildly encouraging. Exports have a positive effect on PLP, most likely because it encourages skilled labor. Disappointing is that increasing FDI has no effect on PLP. This suggests that the increase in low-skill labor by FDI is not enough to offset the effects of a large surplus labor population, even if the actual number of surplus workers is decreasing. The earlier discussion of PLP shows that gains in PLP may not always occur when low-skill labor increases and surplus labor decreases. Based on Olson's logic, even if FDI encourages the employment of low-skilled workers, this group is still hard to organize, particularly if the level of surplus is still high.

Breakdown of Results for HICs

The encouraging findings for skill ratios, surplus labor, and PLP in HICs confirm the primary argument in this chapter that national income level matters. Contrasting the LICs and MICs, portfolio flows, more than

exports and FDI, affect labor. These findings are consistent with research, revealing that private capital flows are more efficient in higher-income countries.[28] In addition, these economies are specializing in services and moving away from manufacturing (ILO, 1997). The data reveal that, compared to the other LDCs, exports of manufactured goods show a distinct decline in HICs since the mid-1980s while GDP growth has increased.

Overall these results indicate that the impact of portfolio flows is beneficial for workers in HICs. Portfolio flows and the savings and investment they generate, more than exports or FDI, encourage the employment of high-skilled labor. As expected, none of the globalization variables affect surplus labor because this population is already small in HICs. Finally, both FDI and portfolio flows lead to greater PLP.

7 Implications

Previous research has not attempted to conduct a comprehensive analysis on the factors affecting workers in the developing world in this era of globalization. We can hope that this study will help inform the views of researchers, politicians, and policy commentators about the effects of globalization on the overall situation of workers in LDCs. Contemporary debates have been polarized. One view is that globalization improves labor's conditions in the developing world because the production and exports of labor-intensive goods will increase, and employment will rise. The opposing view of many comparative and regional specialists is that globalization worsens the bargaining propensity of labor because LDC governments face increasing pressure to keep wages low, at all costs. This analysis suggests the possibility that globalization is improving the economic situation of workers in all developing countries, but it is not universally improving their marketplace bargaining power.

Results from this study challenge broad generalizations about the effects of globalization on labor in LDCs. Econometric analysis on 59 LDCs from 1972–1997 reveals that the effects of globalization on workers are ultimately contingent upon the level of LDC economic development. This analysis serves as a corrective to studies that neglect the importance of the level of economic development in considering the

[28] For a recent review of the literature that supports the hypothesis that private capital flows are more efficient in higher-income countries, see Eichengreen (2001).

domestic effects of globalization. The findings suggest that workers in LDCs at higher levels of economic development are the "winners" in the current era of globalization. Workers in these nations are well positioned to reap the benefits of globalization because of the growing population of skilled labor and the relatively small pools of surplus labor. Workers in middle- and low-income countries do better economically (i.e., greater access to employment opportunities), but are "losers" in the sense that PLP does not significantly increase with globalization.

What are the possible policy implications? This analysis postulates that progressive conditions for workers during globalization rest on their improved bargaining position, and not just pure economic gains. Governments of LICs and MICs must find strategies to reconcile the positive effects of trade on the employment of labor with the expansionary effects it can also have on the size of the informal sector. The most obvious and direct way to do this is to formulate policies that both increase the skill level of the workforce and control the surplus labor population. For instance, governments can provide incentives to firms to hire surplus workers by reducing the extensive labor market regulations that tend to protect privileged labor groups in LDCs. Greater access to training and education will also make workers more employable in the formal sector. Finally, workers themselves may find it more politically beneficial to form coalitions with informal labor groups, sacrificing short run gains (e.g., higher wages) for long-run benefits (greater PLP). Of course, political constraints such as personalistic dictatorships and even partisan politics can thwart the implementation of all these solutions, and the next step in this analysis is to explore such potential obstacles in greater detail. It is hoped that this chapter's findings raise important questions and provide fertile ground for further studies on labor and globalization in LDCs.

Appendix 1

Table 11 Data sources and definitions

*Concepts**	*Measurements*	*Definition*
Globalization [XMANUF] [KFLOW] [FD1]	The amount of total trade (exports + imports/GDP); and gross capital flows as a percentage of GDP; foreign direct investment flows as a percentage of GDP	Gross capital flows are the sum of all inflows and outflows, using the finest classifications to avoid netting
Labor power [PLP] [SKILL] [SURPLUS]	[The ratio of the numbers employed in skill—intensive manufacturing industries relative to numbers employed m low—skill manufacturing industries] times [1 divided by the number of surplus laborers in the economy]/ divided by 87	"Low-skilled" refers to those who have no more than a primary or secondary education, and are likely to be employed in labor-intensive manufacturing industries. "Skilled" are those with more than a basic general education and are usually employed in heavy and high-skill manufacturing industries. "Surplus labor" is the total working age population (between 15 and 65) minus the total labor force minus students enrolled in secondary and tertiary education This total is taken as a percentage of the economically active population. Note that in some poor countries, surplus labor took a negative value most likely because the official working age is much lower than 15 in LDCs. To address this, the number of children ages 10–14 is added to the working age population
Demographic variables [URBAN] [DEPEND]	Urban population as a percentage of total population, the age dependency ratio	"Urbanization" is the midyear population of areas defined as urban in each country. It is measured here as the percentage of the total population "Age Dependency ratio" is the number of persons over 60 divided by number of persons aged 20 to 59
Economic development [GDP] [GROWTH]	The gross domestic product per capita [GDP], GDP growth[GROWTH]	"GDP" is the total gross domestic product of a country divided by total population. "Growth" is the annual percentage growth rate of GDP at market prices based on constant 1987 local currency. Aggregates are based on constant 1987 US dollars

(*continued*)

Table 11 (continued)

*Concepts**	*Measurements*	*Definition*
Political development [DEMOC]	Indicator of democracy	Using scale 0–10, 10=strong democracy. This indicator is derived from the competitiveness of political participation codings, the openness and competitiveness of executive recruitment, and constraints on the chief executive
Socioeconomic development [HUMAN CAPITAL SPENDING]	Government spending on education and health per capita	Include public expenditures in health care and education measured in 1995 constant US dollars

* <.15

Table 12 Data sources, calculations, and estimations

HUMAN CAP:	IMF, **Government Finance Statistics.**
K FLOWS:	IMF, **Balance of Payments Statistics.** World Bank, **World Development Indicators.**
XMANUF:	World Bank, **World Development Indicators.**
SKILL:	Classification scheme developed by Wood and Mayer. Their export product classifications are based on the Standard International Trade Classification (SITC, Revision 2). The following list shows which International Standard Industrial Classification (ISIC) codes correspond to the SITC codes (also supplied by Wood). The employment statistics for each manufacturing sector came from **UNIDO Database of Industrial Statistics.** The final value of SKILL is based on the total numbers employed in high-skill manufacturing production/low-skill manufacturing production.
SURP:	World Bank, **World Development Indicators.**
GDP:	World Bank, **World Development Indicators.**
DEPEND:	World Bank, **World Development Indicators.**
GROWTH:	World Bank, **World Development Indicators.**
DEMOC:	Ted Robert Gurr's and Keith Jaggar's **Polity IV.**

Table 13 Manufactured exports (NM)

Low-skill manufactures	*SITC2 categories*
(1) Leather and rubber products	61–62
(2) Wood and paper products	63–64
(3) Textiles, clothing, footwear, and travel goods	65, 83–85
(4) Nonmetallic mineral products	66 (less 667)
(5) Iron and steel and metal products	67, 69
(6) Furniture and plumbing equipment	81–82
(7) Ships, bicycles, and trains	78 (less 781–4), 79 (less 792)
(8) Miscellaneous	89, 9 (less 941,971)
High-skill manufactures	
(9) Chemicals	5 (less 522.24, 522.56, 524)
(10) Cut diamonds	667.29
(11) Nonelectrical machinery	71–74
(12) Computers and office equipment	75
(13) Communication equipment	76
(14) Electrical machinery	77
(15) Motor vehicles and aircraft	781–784, 792
(16) Scientific instruments, watches, and cameras	87, 88

Note: The SITC 5–8 categories allocated to primary rather than manufactured exports are phosphorus pentoxide and phosphoric acids (522.24), aluminum hydroxide (522.56), radioactive material (524), pearls and precious stones, except cut diamonds (667 except 667.29), and nonferrous metals (**68**)

Table 14 Translated to manufactured production

Low-skill manufacturing production	*ISIC1 categories*
(1) Leather and rubber products	323, 355
(2) Wood products, except furniture Paper and paper products	331, 341
(3) Textiles	321, 323, 322, 324
Leather products	369, 362, 361
Wearing apparel, except footwear	
Footwear, except rubber or plastic	
(4) Other nonmetallic mineral products	
Glass and products	
Pottery, china, and earthenware	
(5) Iron and steel	371, 381
Fabricated metal products	
(6) Furniture, except metal	332
(7) —	
(8) Plastic products	356, 390
Other manufactured products	
High-skill manufacturing production	

(*continued*)

Table 14 (continued)

Low-skill manufacturing production	*ISIC1 categories*
(9) Industrial chemicals Other chemicals Misc. petroleum and coal products Plastic products	351, 352, 354, 356
(10) —	
(11) Fabricated metal products Machinery, except electrical Machinery, electric	381, 382, 383, 384
(12)	—
(13)	—

Appendix 2

Table 15 Comparing PLP and LSI[a]

Countries	*PLP*	*LSI (McGuire)*
Argentina	high	High
Bangladesh	low	Low
Bolivia	med-low	Low
Botswana	low	Low
Brazil	high	High
Cameroon	low	med-high
Chile	med-high	Low
China	high	med-high
Colombia	med-high	med-low
Costa Rica	med-high	med-high
Cyprus	med-high	High
Ecuador	med-low	med-low
Egypt, UAR	med-high	med-high
El Salvador	Low	Low
Ghana	med-high	med-low
Greece	med-high	High
Guatemala	med-high	med-low
Honduras	Low	Low
India	med-low	med-low
Indonesia	med-low	Low
Israel	High	med-high
Kenya	med-high	med-high

(*continued*)

Table 15 (continued)

Countries	*PLP*	*LSI (McGuire)*	
Korea, Rep.	High	low	
Malaysia	High	low	
Mali	med-low	high	
Mauritius	Low	med-low	
Mexico	med-high	med-high	
Morocco	Low	low	
Nicaragua	med-low	med-high	
Nigeria	med-low	med-low	
Pakistan	Low	low	
Panama	Low	low	
Paraguay	med-low	low	
Peru	med-low	med-low	
Philippines	High	med-low	
Singapore	High	low	
South Africa	High	med-low	
Thailand	High	med-low	
Tunisia	med-low	low	
Turkey	med-high	med-high	
Uruguay	High	med-high	
Venezuela	med-high	med-low	
Zambia	med-low	med-high	
Zimbabwe	med-low	med-low	
Percentile Range			
Below 25th percentile	Low	25th percentile-50th percentile	med-low
50th percentile-75th percentile	med-high	above 75h percentile	high

[a]To facilitate direct comparison with McGuire (1999), the PLP values in Table 15 are averages for 1990–1997 only. Several LDCs had to be dropped from the comparison because they were not included in LSI data. Also, because McGuire's index additionally includes developed and Eastern European nations, I eliminated these countries from the sample, recalculated the percentiles, and rechecked the comparison. The results were almost identical to those reported above.

Assessing PLP

Few efforts have been made to measure and compare labor power across developing countries over time. Union density is the most commonly used cross-national indicator of labor power. As noted earlier, union density is more appropriately applied in the developed world than in the LDCs. Most LDCs are still far from attaining strong and independent unions. Even in LDCs with relatively high union density, labor is rife with

collective action problems and often subject to a broad range of government controls.

Given the unreliability of direct organizational measures, as Encarnation's analysis suggests, alternative assessments of labor's bargaining power tend to be tautological (Encarnation, 1989). According to Encarnation, bargaining power is generally defined by the outcome, making it difficult to tell which party had more bargaining power if negotiations are "won by those who win." It is virtually impossible to differentiate between power and negotiated outcomes using this approach. Encarnation (1989, p. 20) concludes that bargaining power must refer to the ability of laborers to "improve the range of plausible outcomes available to each [negotiator], and to improve the probability of securing the outcome that each prefers."

The measure of PLP used in this analysis attempts to avoid the tautology problem. It does so by acquiring some sense of labor's *propensity* for collection action rather than collective action per se. After all, since labor discontent can be costly for political leaders (and workers), governments often respond to labor demands before strikes or other militant actions occur. Offe and Wiesenthal (1985) argue that in such circumstances "the organization then has become strong enough to derive some power (i.e., control over its environment) from its recognized *potential* of power. In other words, concessions are likely to be made not because members have struck, but in order to avoid a strike."

To assess whether PLP serves as an indirect measure of labor's political power, additional steps must be taken. Comparing PLP to other nontautological assessments of labor's bargaining power is the most precise way to accomplish this. McGuire's (1999) labor strength index (LSI) represents the only other effort to assess the "real" magnitude of labor's bargaining power in LDCs and compare it across countries. Because of data limitations, it represents only one period of time (the 1990s). LSI is based on four dimensions: (1) union membership as a percent of the nonagricultural labor force; (2) proportion of formal-sector workers covered by collective contracts; (3) level of collective bargaining power—national/sectoral, enterprise, or both; (4) number of major International Labour Organization conventions ratified.[29] This is a multifaceted attempt to

[29] The number of ILO ratifications is arguably the weakest component of LSI, since ratification does not necessarily ensure enforcement. A detailed evaluation of the strengths and weaknesses of LSI is a subject for a future study.

capture several important dimensions of labor strength that are not directly measured by PLP.

The comparison of PLP and LSI in Table 15 significantly increases confidence in the PLP's reliability as an indicator of labor's bargaining power. The correlation coefficient, excluding the outliers, is 0.61 (see Chart 1). The correlation is actually higher than expected since LSI includes unionization data (and its inherent weaknesses), and because PLP captures some important nontraditional sources of labor's bargaining power.[30]

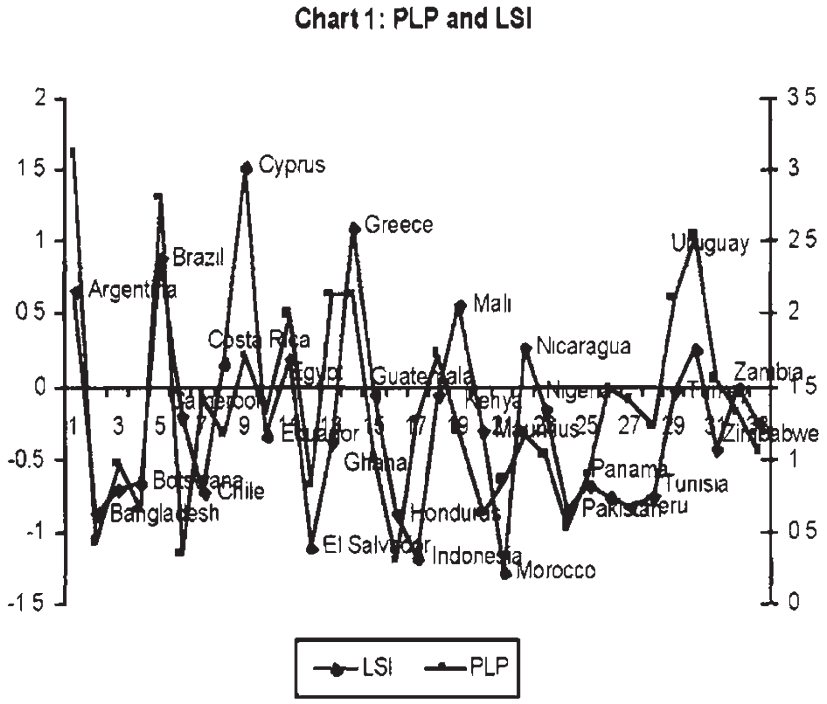

Chart 1 PLP and LSI. (*Note that some of the countries were dropped from the graph to reduce clustering)

[30] LSI closely resembles Silver's (2003, p. 13) reference to "associational power," or "the various forms of power that result from the formation of collective organization of workers" (most important, trade unions and political parties). The comparisons in Table 15 and Chart <InternalRef RefID="Figa">1 suggest that marketplace bargaining power and associational power are closely related.

For most cases, the PLP rankings are similar to LSI, and it is primarily the East Asian cases, for example, Singapore, Korea, Malaysia, and Thailand that show the most contrast. Their PLP score is "high," which contradicts conventional wisdom on labor in these nations.[31] Yet the PLP scores are consistent with more recent analyses by Yap (2003) and Brown (2004), which bring important new insights on labor in the East Asian countries and explain why labor's political influence in these authoritarian nations has been commonly misinterpreted.

According to both Yap and Brown, workers in these nations have been in a relatively unique position in the developing world because of the central economic role they have played in the nations' development. Repressive labor strategies have been part of the East Asian nations' export-oriented industrialization strategy for economic development. Yet paradoxically, precisely because of this dependence on labor, the state has had to accommodate labor in different ways.[32] Labor's collective political consciousness has thereby evolved differently in these countries while the more familiar signs of political power (e.g., strikes, unionization, centralization of bargaining power) have been conspicuously absent.[33] As Young (2004) argues, "in studies … where the forms of consciousness and organization are found not to conform to these [familiar] expectations, labor is deemed to be 'weak' or 'immature,' and seen to be peripheral to the development of state, society and the economy…. The outcome may not conform to

[31] LSI and PLP also differ in some of the African cases where LSI tends to be higher than PLP (Ghana and Mali). The reason is because data availability for all four components of LSI is apt to be scarce in these countries, biasing these scores upward. See McGuire (1999, p. 12).

[32] Yap (2003) for instance, discusses "credible apologies" that East Asian governments make to labor. They may dismiss, demote, or replace certain government officials deemed responsible for the policies that "hurt" labor, downsize or eliminate the relevant agency, or offer reparations. Representatives from academia, labor, or business also may be invited in to review, evaluate, or oversee changes to government.

[33] For example. Yap (2003) draws from Bates (1981) and argues that labor can withdraw economic resources (e.g., alter their production mix, engage in the black market) to protest the government's economic policies. In reference to workers in Thailand, Brown (2004) discusses the importance of taking account of industrial workers and their organization as *potential political actors*. He argues that "even when labor is invisible, in the sense of not being a public, organized actor overtly engaged in formal political processes, the politics of the working class is nonetheless there and is significant. For, behind the scenes, there has been a continual jockeying to channel and control workers and their struggles. This is to ensure that they either do not emerge as a public, organized force, or if they do, they are organized in a manner that is in keeping with the broader economic, ideological, and political interests of those dominating contests for state power" (Brown, 2004, p. 133).

very generalized theoretical expectations, but that calls for re-evaluation and refinement of theory, rather than a dismissal of the significance of working class struggles." One important advancement of the PLP indicator is that it can approximate labor movements that do not develop the familiar institutional forms.

The indicator applied in this study, PLP, offers three broad advantages: (1) it corresponds to conditions specific to the bargaining power of labor in developing countries; (2) it is comparable across LDCs; and (3) it has a time-series component that can capture the dynamic aspects of bargaining power. The first advantage is important because the logic of PLP is based on the particular circumstances faced by labor in LDCs. Desirability of the second two characteristics is more obvious. A standardized measure available over time and across countries greatly reduces the biases that can affect empirical analyses of labor in the developing world.

References

Achen, C. (2000). Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables. Paper Presented at the Annual Meeting of the Political Methodology Section of the American Political Science Association, UCLA, 20–22 July.

Baer, W., & Herve, M. E. A. (1966). Employment and Industrialization in Developing Countries. *The Quarterly Journal of Economics, 80*(1), 88–107.

Bangasser, P. (2000). The ILO and the Informal Sector: An Institutional History. ILO Employment Paper 2000/9.

Banuri, T., & Amadeo, E. (1991). Worlds within the Third World: Labour Market Institutions in Asia and Latin America. In T. Banuri (Ed.), *Economic Liberalization: No Panacea*. Clarendon Press.

Bates, R. (1981). *Markets and States in Tropical Africa*. University of California Press.

Beck, N., & Katz, J. (1995). What to Do and Not to Do with Time-Series Cross-Section Data. *American Political Science Review, 89*(3), 634–647.

Beeson, M., & Hadiz, V. (1998). Labor and the Politics of Structural Adjustment in Australia and Indonesia. *Journal of Contemporary Asia, 28*(3), 291–309.

Berman, E., & Machin, S. (2000). Skill-biased Technology Transfer around the World. *Oxford Review of Economic Policy, 16*(3), 12–22.

Brown, A. (2004). *Labour. Politics, and the State in Industrializing Thailand*. Routledge, Curzon Taylor and Francis Group.

Brunner, A. (2003). The Long-Run Effects of Trade on Income and Income Growth. IMF Institute Series Working Paper No. 03 37.

Candland, C., & Sil, R. (2001). *The Politics of Labor in a Global Age. Continuity and Change in Late-Industrializing and Post-Socialist Economies.* Oxford University Press.

Chan, A., & Senser, R. (1997). China's Troubled Workers. *Foreign Affairs, 76*(2), 104–117.

Dean, J. M. (1995). From Protectionism to Free Trade Fever? Recent Reforms in Developing Countries. *Open Economies Review, 6*(4), 369–385.

Deshpande, S., & Deshpande, L. (1998, May). Impact of Liberalization on Labour Market in India. *Economic and Political Weekly, 30*, L31–L39.

Devarajan, S., Ghanem, H., & Thierfelder, K. (1997). Economic Reform and Labor Unions: A General Equilibrium Analysis Applied to Bangladesh and Indonesia. *World Bank Economic Review, 11*(1), 145–170.

Deyo, F. (1989). *Beneath the Miracle: Labor Subordination in the New Asian Industrialism.* University of California Press.

Eichengreen, B. (2001). Capital Account Liberalization: What Do Cross-Country Studies Tell Us? *World Bank Economic Review, 15*(3), 341–365.

Encarnation, D. (*1989*). *Dislodging Multinationals: India s Strategy in Comparative Perspective.* Cornell University Press.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism.* Princeton University Press.

Frenkel, S. (1993). *Organized Labor in the Asia-Pacific Region.* 1LR Press.

Frenkel, S., & Peetz, D. (1998). Globalization and Industrial Relations in East Asia: A Three-Country Comparison. *Industrial Relations, 37*(3), 282–310.

Frundt, H. (2002). Central American Unions in the Era of Globalization. *Latin American Research Review, 37*(3), 7–53.

Galenson, W. (1962). *Labor in Developing Economies.* University of California Press.

Galenson, W. (1994). *Trade Union Growth and Decline.* Praeger Publishers.

Gallin, D. (2001). Propositions on Trade Unions and Informal Employment in Times of Globalisation. *Antipode, 33*(3), 531–549.

Gereffi, G. (1995). Global Production Systems and Third World Development. In B. Stallings (Ed.), *Global Change, Regional Response* (pp. 100–142). Cambridge University Press.

Ghose, A. (2000). Trade Liberalization and Manufacturing Employment. ILO Working Employment Paper 2000/3, Geneva.

Golden, M., Wallerstein, M., & Lange, P. (1999). Postwar Trade-Union Organization and Industrial Relations in Twelve Countries. In H. Kitschelt, P. Lange, G. Marks, & J. Stephens (Eds.), *Continuity and Change in Contemporary Capitalism.* Cambridge University Press.

Harriss-White, B. (1999). State, Market, Collective, and Household Action in India's Social Sector. In B. Harriss-White & S. Subramanian (Eds.), *Illfare in India: Essays on India's Social Sector in Honour of S. Guhan.* Sage Publications.

Ingerson, A. (1984). The Textile Industry and Working Class Culture in Labor. In C. Bergquist (Ed.), *The Capitalist World Economy* (Vol. 7) in *Political Economy of World Systems Annuals*. Sage Publications.

International Labor Organization. (1997). *World Employment Report, 1996–1997*. ILO.

International Labor Organization. (1999). *World Employment Report, 1998–1999*. ILO.

International Labour Office. (2002). *Key Indicators of the Labour Market, 2001–2002*. Routledge.

Krugman, P., & Venables, A. (1995). Globalization and the Inequality of Nations. *Quarterly Journal of Economics, 110*(4), 857–880.

Kuchta-Helbling, C. (2000). *Barriers to Participation: The Informal Sector in Emerging Democracies*. The Center for International Private Enterprise. http://www.cipe.org/topics/informalsector/papers/, http://www.cipe.org/topics/mformalsector/papers/

Lall, S. (1998). Exports of Manufactures by Developing Countries: Emerging Patterns of Trade and Location. *Oxford Review of Economic Policy, 14*(2), 54–74.

Learner, E. (1996). Wage Inequality from International Competition and Technological Change: Theory and Country Experience. *The American Economic Review, 86*(2), 309–314.

Lewbel, A. (1997). Constructing Instruments for Regressions with Measurement Error When No Additional Data Are Available, with an Application to Patents and R&D. *Econometrica, 65*(5), 1201–1213.

Loayza, N. (1997). The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America. World Bank Working Paper no. 1727.

Lok, H. P. (1993). Labour in the Garment Industry: An Employer's Perspective. In C. Manning & J. Hardjono (Eds.), *Indonesia Assessment 1993*. Australian National University.

Manning, C. (1998). Does Globalization Undermine Labour Standards? Lessons from East Asia. *Australian Journal of International Affairs, 52*(2), 133–147.

Mazumdar, J., & Quispe-Magnoli, M. (2002). Trade and the Skill Premium in Developing Countries: The Role of Intermediate Goods and Some Evidence from Peru. Federal Reserve Bank of Atlanta Working Paper, 2002–11.

McGuire, J. (1997). *Peronism without Peron*. Stanford University Press.

McGuire, J. (1999). Labor Union Strength and Human Development m East Asia and Latin America. *Studies in Comparative International Development, 33*(4), 2–33.

Montiel, P. (1994). Capital Mobility m Developing Countries: Some Measurement Issues and Empirical Estimates. *World Bank Economic Review, 8*(3), 311–350.

Munck, R. (2002). *Globalization and Labour*. Zed Books.

Offe, C., & Wiesenthal, H. (1985). Two Logics of Collective Action. In J. Keane (Ed.), *Disorganized Capitalism* (p. 216). MIT University Press.

Olson, M. (1971). *Logic of Collective Action*. Harvard University Press.
Papola, T. S. (1994). Employment Growth and Social Protection of Labor in India. *Indian Journal of Industrial Relations, 30*(2), 117–141.
Portes, A. (1990). When More Can Be Less: Labor Standards, Development, and the Informal Economy. In *Labor Standards and Development in the Global Economy,* Symposium on Labor Standards and Development U S. Department of Labor: Bureau of International Labor Affairs.
Posusney, M. P. (1997). *Labor and the State in Egypt*. Columbia University Press.
Prasad, S. (1998). Economic Globalization and Labour: Market Transformations in Fiji. *Economic and Industrial Democracy, 19*(2), 227–251.
Robbins, D., & Gindling, T. (1997). Educational Expansion, Trade Liberalization, and Distribution in Costa Rica. In A. Berry (Ed.), *Poverty, Economic Reform, and Income Distribution in Latin America*. Lynne Rienner Publishers.
Rogowski, R. (1989). *Commerce and Coalitions*. Princeton University Press.
Scruggs, L., & Lange, P. (2002). Where Have All the Members Gone? Globalization, Institutions, and Union Density. *The Journal of Politics, 64*(1), 126–153.
Sen, A. (1966). Peasants and Dualism with or without Surplus Labor. *The Journal of Political Economy, 74*(5), 425–150.
Silver, B. (2003). *Forces of Labor*. Cambridge University Press.
Stone, A., Levy, B., & Paredes, R. (1996). Public Institutions and Private Transactions: A Comparative Analysis of the Legal and Regulatory Environment for Business Transactions in Brazil and Chile. In L. Alston (Ed.), *Empirical Studies in Institutional Change*. Cambridge University Press.
United Nations. (1993). *UN Industry and Development Report. 1992/1993*. United Nations Industrial Development Organization.
Valenzuela, S. (1989). Labor Movements in Transitions to Democracy. *Comparative Politics, 21*(4), 445–172.
Webster, E., & Lipsig-Mumme, C. (2002). Recasting Labour Studies in the New Millennium. *Society in Transition, 33*(2), 258–265.
Wellisz, S. (1968). Dual Economies, Disguised Unemployment and the Unlimited Supply of Labour. *Economica, 137*, 22–51, New Series 35.
Western, B. (1997). *Between Class and Market. Postwar Unionization in the Capitalist Democracies*. Princeton University Press.
Wood, A. (1997). Openness and Wage Inequality in Developing Countries: The Latin American Challenge to the East Asian Conventional Wisdom. *World Bank Economic Review, 11*(1), 33–57.
Wood, A., & Ridao-Cano, C. (1999). Skill, Trade, and International Inequality. *Oxford Economic Papers, 51*, 89–119.
World Bank. (1995). *World Development Report 1994–1995*. World Bank.
World Bank. (2000). *World Development Report 1999–2000*. World Bank.
World Bank. (2001). *World Development Report 2000–2001*. World Bank.

World Bank. (2002). *World Development Report, 2002*. World Bank.
Wright, E. (2000, January). Working-Class Power, Capitalist-Class Interests, and Class. *American Journal of Sociology, 105*, 957–1002.
Yap, O. F. (2003). Non-Electoral Responsiveness Mechanisms: Lessons from the Asian Less-Democratic Newly Industrializing Countries. *The British Journal of Political Science, 33*, 491–514.
Young, K. (2004). Review of Labour, Politics, and the State in Industrializing Thailand. *Journal of Contemporary Asia, 34*(4), 551–552.
Yusuf, S. (1999). The Changing Development Landscape. *Finance and Development, 36*(4), 15–18.

The Role of Labour in Capability Upgrading: The Case of Emerging Market Multinationals

Vito Amendolagine and Roberta Rabellotti

1 Strategic Asset-Seeking Investments and Emerging Market Multinationals

Since the beginning of 2000s, foreign direct investments (FDIs) by Emerging Market Multinationals (EMNEs) have increased their shares on worldwide FDIs (Fig. 1). With a focus on BRICS (Brazil, Russia, India, China and South Africa) starting from just 0.8% in 2001, their FDIs have reached 17% of worldwide FDI outflows in 2020 (UNCTAD, 2021). Notwithstanding recent concerns on foreign ownership fuelled by national security considerations (UNCTAD, 2019), emerging countries still rank in top positions as both FDI origins and destinations. In 2020 China is the second host country of FDIs in the world (149 US$ billion), India the

V. Amendolagine
Dipartimento di Economia, Università di Foggia, Foggia, Italy
e-mail: vito.amendolagine@unifg.it

R. Rabellotti (✉)
Dipartimento di Scienze Politiche e Sociali, Università di Pavia, Pavia, Italy
e-mail: roberta.rabellotti@unipv.it

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_14

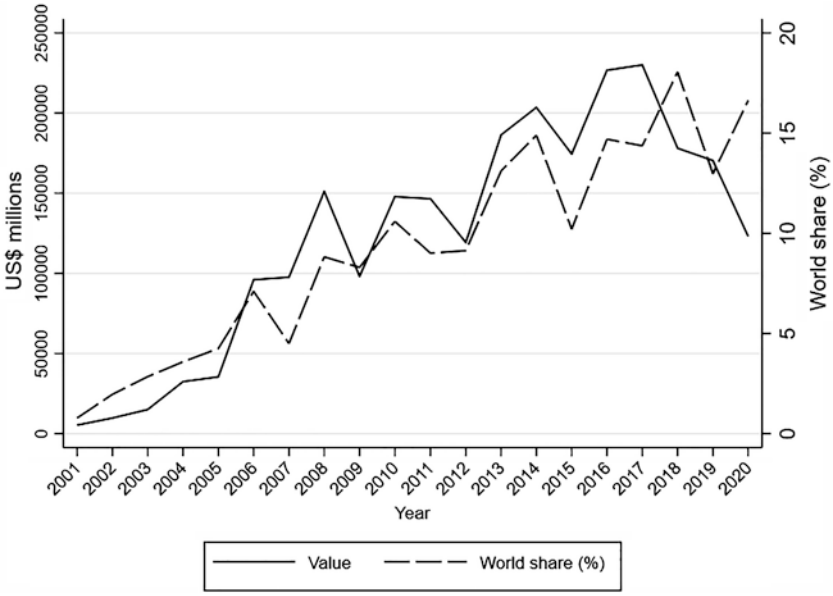

Fig. 1 Foreign direct investments (FDIs) by Emerging Market Multinationals (EMNEs)—shares on worldwide FDIs. (Source: UNCTAD (2021))

fifth (64 US$ billion) and Brazil the eleventh (25 US$ billion). Considering outflows, China is the third most important FDI origin country in the world (133 US$ billion) and India the twentieth (12 US$ billion) (UNCTAD, 2021).

According to UNCTAD (2017), around 40% of FDIs from emerging countries target developed countries with two main motivations: strategic assets and market seeking (Ramamurti, 2012). The literature agrees that EMNEs lack the competitive advantage to compete local firms in foreign markets and therefore they seek to augment their strategic assets through the acquisition of a variety of intangible resources such as technologies, R&D capability, managerial competency, human capital, brand names and commercial knowledge (Buckley et al., 2016).

EMNEs gain a quick access to advanced knowledge and technological assets directly through acquisitions of technological leading companies. Alternatively, they undertake greenfield investments in technologically

advanced hubs, connecting with local actors such as suppliers, universities and service providers (Meyer et al., 2011).

But how does the knowledge acquisition and capabilities upgrading process take place in EMNEs undertaking strategic asset-seeking investments in developed countries? A key role is indeed played by the access to specialized workers, such as scientists, managers and engineers operating either in the EMNEs' subsidiaries or through tapping into local networks including other local firms and institutions such as R&D labs or universities. In the rest of this note we provide an, incomplete, account of some interesting contributions which explore how access to skilled labour can contribute to capability upgrading in emerging multinationals, also discussing some of the challenges EMNEs often face.

2 The Key Role of Skilled Labour for Acquiring Knowledge

The role played by staff mobility between host and home countries has been studied in several case studies mainly with a focus on Chinese multinationals. The establishment of an independent R&D company—China Euro Vehicle Technology (CEVT)—by Geely Auto and Volvo is investigated by Yakob et al. (2018), who stress that the intensive collaboration between Sweden and China is mainly based on exchanges of engineers and other managers. These authors emphasize that learning is a two-way process with Chinese engineers visiting Sweden to improve their technological competences and Swedish managers learning from Chinese employees about the booming Chinese automotive market.

By exploring the operations in developed countries of two Chinese ITC leading multinationals—Huawey and ZTE—Fu et al. (2018) show how the integration of the headquarters with their subsidiaries' human resources, including engineers, managers and technicians, is key for the upgrading of investors' technological capabilities. They argue that such integration of resources takes place through three different but related steps: (i) leveraging subsidiaries' resources; (ii) generating new resources; and (iii) releasing resources. The authors explain that to develop new products and business capabilities, headquarters and subsidiaries combine their knowledge by exchanging not only production components but also specialized personnel; however, in some cases innovating can also imply

releasing some of the existing resources, by selling assets or substituting part of the existing workforce.

In addition to staff mobility within multinationals, the access to the specialized labour market of the local context where EMNEs invest also matters for their learning opportunities. The seminal work by Almeida and Kogut (1999), focusing on the US semiconductor industry, shows that regions are significantly different in terms of knowledge endowments and this happens because they differ in the specialized labour markets for scientists, engineers and other skilled workers. Thus, innovative hubs such as the Silicon Valley enjoy strong technological spillovers, triggered by intense local inter-firm mobility of skilled workers. These spillovers are also important for emerging market multinationals seeking to acquire knowledge in developed countries. In the case of Chinese and Indian cross-border acquisitions in EU-28 and USA, Amendolagine et al. (2018) show that knowledge transfer, also occurring through skilled labour mobility, takes place not only within the company, between the subsidiaries and the headquarter, via expatriate managers, but also within the regional context where the subsidiaries are located. EMNEs making investments in regions with strong innovative capacity have opportunities to tap into the regional pool of specialized labour leveraging the acquired firms' linkages to other actors in the regional ecosystem.

Liu and Giroud (2016) highlight the role of high-skilled returnee migrants as another mechanism that can help integrating investors' and subsidiaries' innovative resources. By studying or working abroad, returnees accumulate knowledge and experience about international markets that can support EMNEs when sourcing technology from developed markets. In some cases, for EMNEs employing returnee managers or CEOs can compensate for the liability of emergingness, that is, a country of origin disadvantage, and for their lack of international experience. Particularly, high-skilled returnee employees' language and cultural competencies can enhance trust in acquired companies and in the local host ecosystem, facilitating knowledge sharing and learning and overcoming some of the challenges faced by EMNES, described in the next section.

3 Labour Integration Between EMNEs and Developed Countries' Subsidiaries Is Not Always an Easy Task

EMNEs are not easily able to take advantage of their FDIs in developed countries to improve their innovative capabilities, acquiring knowledge and learning from their subsidiaries and from other key actors with whom they interact in the host regions. Amendolagine et al. (2018) find that Chinese and Indian acquisitions in developed countries do not always boost investors' innovative capacity and identify two moderating factors influencing their capabilities to learn and acquire knowledge: (a) the *absorptive capacity*, which is very heterogenous among investors; and (b) the *status*, that is, the perception of EMNEs' qualities, which affect the willingness of the local skilled workers and other relevant actors to share knowledge. Due to the investors' liability of emergingness, managers and other skilled workers (such as researchers, engineers and technicians) in acquired companies might be unwilling to collaborate with the new owners. In some cases, the most skilled and talented personnel might even leave the company, so reducing the knowledge endowment available within the multinational boundaries. Similar considerations are likely to apply to the host ecosystem where the local actors might fear loss of their knowledge, being unwilling to collaborate with EMNEs. Therefore, only an elite of EMNEs characterized by a strong absorptive capacity and high reputational status is able to effectively benefit from knowledge transfer through the access to local high-skilled human resources both in the acquired companies and in the host regions.

The importance of the perceived reputation of EMNEs is confirmed by an empirical study by Rao-Nicholson et al. (2016), inspired by the lively debate in the media after the acquisition of the British steel producer Corus by the Indian multinational Tata, in which the main narrative was about the possible loss of local jobs as well as other issues related to pension package that UK employees were expected to receive. The authors investigate the impact of Chinese and Indian cross-border acquisitions on the psychological safety of employees in the acquired companies, using survey methods to explore their expectations in terms of job stability and remuneration after the takeover. They find that the employees in the acquired firms experience negative emotions when the leadership of EMNEs is very active (i.e. proactive engagements of top managers in deal-related activities or in meetings with workers and trade unions). Local

employees and managers feel insecure and fear a loss of autonomy in the target company (particularly in developed countries). When instead there is trust in leadership ability, integrity and reputation, the psychological climate within the acquired company improves and workers feel safer and more willing to share their knowledge, sustained by the evidence of value congruence between the acquirer and the local employees.

4 Conclusions and Implications for Further Research

This note shows how the access to skilled labour represents a key mechanism for acquiring knowledge and learning in EMNEs undertaking strategic asset-seeking investments in developed countries. It indicates that there are two main channels: (a) within the multinational the staff mobility between host and home countries; and (b) within the host local context the access to specialized skilled labour market. The access to knowledge through these channels cannot be taken for granted and only EMNEs with strong absorptive capacity and high reputational status are able to overcome the liability of emergingness which can hamper their process of building up innovative capabilities.

In terms of further research, it would be highly relevant to collect more micro empirical evidence about the learning mechanisms from staff mobility and integration in locally embedded knowledge networks and about the range of frictions EMNEs face in the process of integration. Further questions to be addressed by future empirical research concern possible facilitators which can help less experienced EMNEs to access local human resources and networking with local partners as well as the investigation of the pathways for reverse knowledge to understand how new knowledge is disseminated and integrated within EMNEs. Finally, a key issue requiring more research is how different entry modes influence routines of knowledge accumulation and human resources integration in EMNEs and in the host local economies.

References

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science, 45*(7), 905–917.

Amendolagine, V., Giuliani, E., Martinelli, A., & Rabellotti, R. (2018). Chinese and Indian MNEs' Shopping Spree in Advanced Countries. How Good is it for Their Innovative Output? *Journal of Economic Geography, 18*(5), 1149–1176.

Buckley, P. J., Munjal, S., Enderwick, P., & Forsans, N. (2016). Do Foreign Resources Assist or impede Internationalisation? Evidence from Internationalisation of Indian Multinational Enterprises. *International Business Review, 25*(1), 130–140.

Fu, X., Sun, Z., & Ghauri, P. N. (2018). Reverse Knowledge Acquisition in Emerging Market MNEs: The Experiences of Huawei and ZTE. *Journal of Business Research, 93*, 202–215.

Liu, X., & Giroud, A. (2016). International Knowledge Flows in the Context of Emerging-economy MNEs and Increasing Global Mobility. *International Business Review, 25*(1), 125–129.

Meyer, K. E., Mudambi, R., & Narula, R. (2011). Multinational Enterprises and Local Contexts: The Opportunities and Challenges of Multiple Embeddedness. *Journal of Management Studies, 48*, 235–252.

Ramamurti, R. (2012). What is Really Different about Emerging Market Multinationals? *Global Strategy Journal, 2*, 41–47.

Rao-Nicholson, R., Khan, Z., & Stokes, P. (2016). Making Great Minds Think Alike: Emerging Market Multinational Firms' Leadership Effects on Targets' Employee Psychological Safety after Cross-border Mergers and Acquisitions. *International Business Review, 25*(1), 103–113.

UNCTAD. (2017). *World Investment Report*. UNCTAD, Geneva.

UNCTAD. (2019). *World Investment Report*. UNCTAD, Geneva.

UNCTAD. (2021). *World Investment Report*. UNCTAD, Geneva.

Yakob, R., Nakamura, H. R., & Ström, P. (2018). Chinese Foreign Acquisitions Aimed for Strategic Asset-creation and Innovation Upgrading: The Case of Geely and Volvo Cars. *Technovation, 70*, 59–72.

Employment Effect of Innovation

d'Artis Kancs and Boriss Siliverstovs

1 Introduction

In setting the Research and Innovation (R&I) Strategy 2020–2024, the European Union (EU) has defined six ambitious objectives; including jobs and employment. Research and innovation policy should play a key role in responding to the challenges brought about by the global COVID-19 pandemic. It should help deliver Europe's recovery plan, paving the way out of the current crisis on the path to a fairer future, based on economic growth that places the wellbeing of workers at the centre of the production process (Ivanova et al., 2019).

In the context of these R&I Strategy's objectives, an important policy question arises whether innovation and employment processes can be

d'A. Kancs (✉)
Joint Research Centre, European Commission, Ispra, Italy

LICOS, University of Leuven, Leuven, Belgium
e-mail: d'artis.kancs@ec.europa.eu

B. Siliverstovs
Bank of Latvia, Riga, Latvia

KOF Swiss Economic Institute, ETH Zurich, Zurich, Switzerland

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_15

complementary and hence their EU targets can be achieved at the same time? Further, policy makers are interested to know: (i) are there R&D intensity levels when innovation and employment are positively related to each other and when innovation may have an adverse impact on the firm employment? (ii) What type of innovators create most jobs and hence provide the highest potential for policy synergies? Answering these questions is the main objective of this study, as they may help to design policies, which can efficiently contribute to achieving both the innovation and employment targets of the R&I Strategy at the same time.

At a first glance, a simultaneous boosting of both employment and innovation may seem an easy and most natural task to achieve as any type of investments (including R&D) increases the labour demand, at least in the short-run. However, the theoretical literature suggests that the relationship between innovation and employment seems to be far more complicated than one can naively assume initially (Smolny, 1998). Also the econometric results reported in the literature on employment effects of innovation are rather contradictory both with respect to their sign and magnitude, suggesting that increasing the innovation intensity can have not only complementary but also substitutionary effects on employment (Young, 1993; Antonucci & Pianta, 2002; Van Reenen, 1997).

In order to accommodate a wide range of possibilities in the innovation-employment relationship ranging from highly negative to strongly positive, in this study we propose an alternative methodological approach that has not been employed in the innovation-employment literature before. In particular, we relax the linearity assumption in the functional relationship between innovation and employment and hope that it will contribute towards sorting out the likely reasons for observing such a large range of estimated employment elasticities with respect to the firm innovation activity. There are several reasons why the innovation-employment relationship may be non-linear. Conceptually, the non-linearities in the functional relationship between innovation and employment may arise, for example, due to the coexistence of many mutually interdependent transmission mechanisms and general equilibrium feedback loops, as the employment effect of innovation depends, among others, on the nature of innovation (product or process innovation); the purpose of innovation (to save labour or capital, neutral, or biased towards skills) and other factors (Pianta, 2004). Empirically, the employment effect of innovation depends on the firm's sector of activity; formal and informal institutions; the time frame of analysis; specifics of the existing production technology;

dimensions of innovation (radical or incremental); consumer preferences; the fierceness of competition in intermediate input and labour markets; the structure of workforce skills; and so on which all contribute to differentiated employment effects at different innovation intensities (Bogliacino & Vivarelli, 2012; Bogliacino et al., 2012; Lachenmaier & Rottmann, 2007).

If the functional relationship between innovation and employment would indeed be non-linear—a fact confirmed in our econometric analysis—then an accurate estimation of the functional relationship would depend crucially on the ability to account for these non-linearities in the innovation-employment nexus, which is highly challenging. Due to complexities related to a suitable counterfactual at the firm level and methodological challenges in the estimation approach, however, there are no studies available in the literature yet that would attempt to account for non-linearities in the R&D and firm employment relationship in a continuous non-linear setting. This study attempts to fill this research gap and estimate the full functional relationship between the firm's innovation and employment in a continuous setup.

To achieve this objective, we rely on a flexible semi-parametric method—the generalised propensity score (GPS) estimator—suggested by Hirano and Imbens (2004). Two main features of the GPS methodology make it particularly attractive for our purpose: (i) estimation can be based on a flexible semi-parametric regression allowing for a non-linear dependence between the variables of interest without imposing any a priori restrictions; and (ii) the elimination of the selection bias arising from a non-random assignment of treatment (R&D expenditure) intensity across firms by conditioning on the observed firm characteristics. In applying the GPS methodology, we attempt to identify the R&D intensity levels under which innovation can be complementary to employment and under which it may have an adverse impact on employment. To the best of our knowledge, the application of a flexible semi-parametric counterfactual methods to the employment-innovation nexus is the first of this sort in literature and hence constitutes our main contribution to literature.

We base our micro-econometric analysis on a large international firm-level panel data set for OECD countries and our proxy for technology is a measurable and continuous variable, while most of previous studies have relied on either indirect proxies of the technological change or dummy variables (such as the occurrence of product and process innovation). In particular, we employ the EU Industrial R&D Investment Scoreboard data set, which comprises data of the R&D investment, as well as other

financial and economic variables for the top 2500 innovators worldwide. In addition to firm-level R&D expenditures, we make use also of other economic and financial variables, which allow us to control for important firm-specific characteristics. Moreover, the Scoreboard data also allow to identify the industrial sector (of the parent subsidiary) as well as the geographical region of the R&D investment (according to the location of the firm's headquarter), which allows us to control for fixed sector-specific and location-specific effects.

Our results enhance previous findings by facilitating to connect dots of existing point estimates in literature. Our findings confirm that the relationship between innovation and employment entails important non-linearities. There is notable difference in reaction of employment to the innovation activity of the firm, depending on the actual level of the R&D intensity. It is also worthwhile mentioning that our results also remind that the innovation impact on employment can be negative too—findings that have been reported also in previous studies (Pianta, 2004). For example, in our sample this is the case for companies operating in high-tech sectors, characterising by comparatively high levels of the innovation activity. These results imply that a further increase in R&D expenditures in high-tech sectors can have a non-negligible labour-saving effect. Furthermore, we find that the labour-saving effect of innovation could also be detected for companies operating in low- and medium-low-tech sectors, though this effect is much less pronounced than for highly innovative firms.

The rest of this paper is organised as follows. Next section contains a review of the relevant literature. In Sect. 3 we describe the econometric methodology. The data is described in Sect. 4. The empirical results are presented in Sect. 5. The final section contains conclusions and sets an outline for future research agenda.

2 Previous Literature

The question of whether the technological change creates or destroys jobs has been posed since the beginning of the classical economics of Karl Marx (1867):

> "*Suppose that the making of the new machinery affords employment to a greater number of mechanics, can that be called compensation to the carpet makers, thrown on the streets?*" (Marx (1867): 479).

Ciriaci et al. (2013), Bogliacino et al. (2012), Bogliacino and Vivarelli (2012) and Bogliacino (2014) were among first attempts to decompose the employment effect of innovation according to R&D intensity levels. Using a balanced panel comprising 3300 Spanish firms observed of the period 2002–2009, Ciriaci et al. (2013) investigated the employment effect of innovation both for innovative and non-innovative firms. Ciriaci et al. (2013) found that those firms, which engage more intensively in innovation activities, create more jobs than less innovative firms. In particular, this effect is more pronounced for small and young innovative firms. At the same time they pointed out that for this group of firms, a successful launch of new products in the market as a result of boosting the innovation activity can lead to a higher growth in sales rather than in employment, which is consistent with the labour-saving effects of technological advances, discussed above.

Bogliacino et al. (2012) studied the employment effect of R&D expenditure using the sample of 677 EU firms observed during the period 1990-2008. Employment elasticities were estimated using a dynamic panel model allowing for lagged employment by means of the Least Squares Dummy Variable Corrected (LSDVC) estimator (Bun & Kiviet, 2003; Bruno, 2005). The results were obtained for the sample of all firms as well as for sub-samples comprising service-sector firms, all manufacturing firms and sub-samples comprising manufacturing firms further subdivided into high-tech and non-high-tech firms. The estimated short-run elasticities were 0.023% for the whole sample, 0.056% for service-sector firms and 0.049% for high-tech manufacturing firms. Interestingly, also the corresponding elasticity estimate for non-high-tech manufacturing firms was also positive (0.021%), though not statistically significant. Using the estimated coefficient on the lagged employment variable Bogliacino et al. (2012, Table 1) derived long-run employment elasticities. The long-run elasticities of employment calculated for the whole sample were 0.075%, 0.097% for service-sector firms and approximately of equal magnitude of 0.11% both for all manufacturing firms and high-tech manufacturing firms.

Bogliacino and Vivarelli (2012) conducted study on the employment effect of innovation activity using a sample of 2295 firms from 15 European countries available over the period 1996–2005. All main results of this study were reported for a number of dynamic panel data estimators such as random-effects, fixed-effects as well as two versions of the Generalised Method of Moments [GMM-DIF, Arellano & Bond (1991)] and

[GMM-SYS, Blundell & Bond (1998)], where the last estimator could be identified as the most reliable one (Bogliacino and Vivarelli, 2012, Section IV). These estimators were applied for the whole sample of firms. The short-run elasticity reported by the GMM-SYS estimator was 0.025%, which was very similar to that reported in Bogliacino et al. (2012). However, the long-run elasticity was about 0.31%, which was about four times larger than that reported in Bogliacino et al. (2012) for the whole sample (0.075%). In order to ensure robustness of estimation results, a distinction was made between firms with different levels of the technological sophistication, by allowing for differential employment effects of high-tech, medium-tech and low-tech firms. Employment elasticities were obtained by means of the LSDVC rather than the GMM estimator; as the former estimator outperformed the latter one under given estimation conditions. The main result of Bogliacino and Vivarelli (2012) was that the job creation effect of the R&D expenditure only was evident for the high-tech sector; both for medium- and low-tech sectors the estimated short-run elasticities were not significantly different from zero. For the high-tech sector, short- and long-run elasticities were 0.017% and 0.17%, respectively.

3 Econometric Strategy

In light of the diversity in the channels of adjustment and the reverse causality of interdependencies between innovation and employment, the existing evidence discussed in Sect. 2 suggests that very likely the functional relationship between these two processes is more nuanced than point estimates from previous studies are able to tell us. This implies that an accurate estimation of the functional relationship depends crucially on the ability to account for potential non-linearities in the innovation-employment nexus. In order to allow for a differentiated impact of innovation on employment while accounting for differences among firms at different R&D intensity levels, an appropriate estimation approach is required which does not average across all innovators and employers, but instead allows for a differentiated employment effect at various R&D intensity levels.

To estimate the full functional relationship between innovation and employment, we rely on the generalised propensity score (GPS) approach introduced in Hirano and Imbens (2004).[1] The GPS approach is a further

[1] This approach was already applied to the following pairs of variables: R&D intensity and productivity in Kancs and Siliverstovs (2016), migration and trade in Egger et al. (2012) and

elaboration on the popular binary treatment propensity score estimator of Rosenbaum and Rubin (1983) widely used for impact evaluations of various programmes.[2] In the context of this study, the relevant features of the GPS methodology are as follows. First, it allows for continuous rather than binary treatment levels. Second, it allows to estimate the treatment effect also without a "zero" control group. Third, the GPS procedure eliminates the selection bias arising due to a non-random assignment (choice) of treatment (R&D) intensity across firms by conditioning on observed firm characteristics. Finally, it captures potential non-linearities in the functional relationship between the R&D investment and firm employment, as it relies on a flexible semi-parametric specification.[3] As result, the estimated dose-response functions allow to retrieve the entire interval of average and marginal treatment effects over all possible treatment levels (R&D intensity).

The counterfactual framework of the dose-response analysis naturally involves a dose or treatment variable—R&D intensity—and a response variable—employment—both observed for firm i. The difference between usual analysis, typically based on the OLS regression of the response variable on the treatment variable, is that one introduces an additional auxiliary variable, called the generalised propensity score, when modelling the dose-response relationship between the variables of interest. The generalised propensity score is derived from a vector of observed covariates for firm i, X_i, and its primary purpose is to remove estimation and inference biases related to non-random dose assignment in the data sample, as discussed above.

Application of the GPS methodology in order to estimate the dose-response analysis typically involves the following three steps (Hirano & Imbens, 2004) . In the first step the GPS variable is constructed using the OLS regression of the treatment variable, r_i, or, as most often in literature, its logarithmic transformation, $\ln r_i$, on a vector of continuous and categorical covariates, X_i, characterising each firm i in the data set:

growth effects of the regional policy in the European Union in Becker et al. (2012), *inter alia*.

[2] For an accessible presentation of the logic underlying the propensity-score matching, see Heinrich et al. (2010).

[3] According to Bia et al. (2011), the estimated dose-response function is robust to the choice of a semi-parametric approach, but it is sensitive to a parametric specification.

$$\ln r_i = X_i'\gamma + \varepsilon_i, \qquad \varepsilon_i \sim N(0,\sigma^2). \tag{1}$$

Observe that a usual assumption that is made is that the distribution of the error terms is normal with variance σ^2. If this assumption is supported by the data then the GPS variable is defined as the normal probability density function estimated for the regression residuals:

$$\hat{s}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left[-\frac{1}{2\hat{\sigma}^2}\hat{\varepsilon}_i^2\right]. \tag{2}$$

Hirano and Imbens (2004, Section 7.4) mention that other more flexible distributions can be used in case if the normality assumption is not supported by the data, for example, this could be a mixture of several normal distributions or a normal heteroscedastic distribution, in which variance is a function of covariates. Alternatively, the departures from normality can be accommodated by non-parametric approach that relies on the kernel probability density estimation. In the empirical part of this paper, we resort to the latter option, since we find that the normality assumption is violated in our sample. To the best of our knowledge, this is the first attempt to rely on non-parametric methods for estimation of the generalised propensity score in the literature and it serves as an additional methodological contribution to the relevant literature.

The propensity score in Eq. (2) fulfils its purpose of measuring the degree of similarity across heterogeneous firms when the so-called *balancing* property is satisfied, that is, for those firms with assigned equal propensity scores (conditional on firm-specific covariates) the associated treatment level is independent of firm characteristics. In this step, we follow the test procedures of Hirano and Imbens (2004) in order to verify whether the balancing property is not violated in our data sample.

In the second step, the expected value of response variable, $\ln \omega_i$, is modelled as a flexible semi-parametric function of the treatment variable and the estimated generalised propensity score, $\ln r_i$ and s_i, respectively:

$$\begin{aligned} E[\ln \omega_i \mid \ln r_i, s_i] = {} & Incpt + \alpha_{11} * \ln r_i + \alpha_{12} * [\ln r_i]^2 + \alpha_{13} * [\ln r_i]^3 \\ & + \alpha_{21} * s_i + \alpha_{22} * [s_i]^2 + \alpha_{23} * [s_i]^3 \\ & + \alpha_3 * (\ln r_i * s_i), \end{aligned} \tag{3}$$

where the latter variable is substituted with its estimates, $\hat{s}_i$, from the first step. The flexibility of the functional form can be controlled for by varying the power of variables $\ln r_i$ and s_i and their cross-products.

The average expected response of the response variable, ω, for a given treatment dose, ρ, is estimated in the third step:

$$E[\ln \hat{\omega}(\ln \rho)] = \frac{1}{N}\sum_{i=1}^{N}\Big[Incpt + \hat{\alpha}_{11} * \ln \rho + \hat{\alpha}_{12} * \left[\ln \rho\right]^2 + \hat{\alpha}_{13} * \left[\ln \rho\right]^3 + \hat{\alpha}_{31} * \hat{s}(\ln \rho, X_i) + \hat{\alpha}_{32} * \left[\hat{s}(\ln \rho, X_i)\right]^2 + \hat{\alpha}_{33} * \left[\hat{s}(\ln \rho, X_i)\right]^3 + \hat{\alpha}_3 * (\ln \rho * \hat{s}(\ln \rho, X_i))\Big], \tag{4}$$

where the coefficient estimates from Eq. (3) are used. The whole dose-response function is obtained by computing Eq. (4) for each treatment level by using a grid of values in the corresponding range of the treatment variable.

In the final step, we derive the treatment effect function as a first derivative of $E\left[\mathrm{I}\hat{\omega}\left(\mathrm{In}\rho\right)\right]$ with respect to argument $\ln \rho$. By definition the treatment effect function computed in this way measures estimated employment elasticity with respect to R&D, allowing us to directly compare our results with those reported in the existing literature. Following Hirano and Imbens (2004), confidence intervals around the estimated dose-response and treatment effect functions are obtained by means of a bootstrap procedure.

4 Data Sources, Sample and Variable Construction

Data Sources

The principal data source is the EU Industrial R&D Investment Scoreboard maintained by the European Commission. The R&D Scoreboard is an annual data set that comprises firm-level data on the R&D investment, as well as other financial and economic variables (e.g. net sales, operating profits, employees) for the top 2500 R&D performers worldwide. In addition to economic and financial variables, the R&D Scoreboard also identifies the main industrial sector (of the parent company) as well as the geographical region of R&D investment (according to the location of company's head-quarter).

An important limitation of the R&D Scoreboard data concerns the issue of non-random sample selection, putting under question the general validity of our results. Given the underlying sampling and selection rules of the R&D Scoreboard data set—ranking and selecting companies according to the total amount of their R&D expenditures—the R&D Scoreboard is not a random sample. Hence the R&D Scoreboard data set may be criticised that it has a sample bias affecting the results, as it only represents top R&D investors. However, given our interest in the employment effect of innovation, this issue is of lower order of magnitude, because we are covering almost the entire population of the world-wide R&D investment (Moncada-Paterno-Castello et al., 2010). As described below, out of the 2500 firms listed in the R&D Scoreboard data only for 1659 companies there were complete data records, prompting us to analyse the available data.[4] Still, these 1659 Scoreboard's companies selected for this study represent around 80% of the world-wide business R&D expenditure. While small R&D investors and non-R&D-performers are excluded from our sample, the aim of this study is to focus on the impact of the R&D-driven innovation on employment, but not to examine determinants of the labour demand in the entire economy. Finally, the particular estimation approach that we adopt in this study allows us to estimate counterfactual treatment effects also without a zero control group.

Sample Construction

In this study, we use R&D Scoreboard data for the last four available years: 2014-2017. Our choice of this sample is motivated by the fact that it is a reasonably long period apart from the Great Financial Crisis (GFC) that undoubtedly had pronounced effects on the firms' investment activity. Including observations from years during the GFC and shortly after its outbreak had a distortive impact on the long-run relationship between innovation and employment prevailing in the business-as-usual environment that we aim to capture in our study.

Since the Scoreboard involves individual firm-level data covering many countries, industries and technological levels of sophistication, it is rather unsurprising that due to all this incumbent data heterogeneity the

[4] Companies which do not disclose figures for R&D investment or which disclose only figures which are not material enough were also omitted from our analysis.

annual data for top 2500 Scoreboard companies forms an unbalanced panel. There are firms that were not present among the top 2500 R&D performers either in the beginning or in the end of the sample period or even at the both ends of the sample period and hence have missing observations. There are also firms that were present in the top 2500 Scoreboard sample at the beginning and at the end but have missing data points for some years within our sample period. All this implies a loss of observations, if our identification strategy aimed at exploiting both inter-temporal and cross-sectional dimensions of a balanced panel. Another option would be to focus solely on the cross-sectional dimension for a particular year, but this again involves loss of information as well as a certain arbitrariness in the choice of the particular year. Hence, in order to retain as many observations as possible, we construct our sample from firms for which there are at least two consecutive years of observations for all variables of our interest. For these firms, we compute averages of their characteristics using the available observations. This helps us smoothing year-on-year fluctuations in our data and avoid a potential source of outlier bias.

Finally, we did a sanity check for the resulting sub-sample of firms and filtered out firms that have extreme values of the R&D intensity which, as discussed in Sect. 4, is defined as the ratio of the R&D investment to net sales. In particular, we removed firms for which the estimated R&D intensity exceeds unity. For this sub-group of firms the median R&D intensity is 6, whereas the maximum is 1210. It turns out that *all* these firms are characterised by a rather small actual employment (the median employment is 113 persons) and a negative operating profit. The former fact indicates that the share and hence the impact of these firms on the total employment are rather small. Moreover, the latter fact indicates that such business model/innovation pattern is not sustainable in the long run. Therefore, in order to make our sample more homogeneous we treat these firms as outliers that need to be removed from the empirical analysis. As a result of data cleaning, we are left with 1659 observations that form the basis for our empirical analysis.

Data Set

The dependent (response) variable is a firm-specific employment measured by the number of employees (EMPL). For each firm in our sample we use the average number of employees for the available years. These

companies included in our sample data employed around 44.1 mln. workers with largest shares of about 10.6 and 14.5 mln. workers pertaining to companies registered in the US and the EU. The R&D investment totalled 2028 milliard Euro with about 42% and 28% of the total sum is attributable to the companies from the US and the EU, with the Japan and China accounting for about 17% and 6%, respectively.[5]

The remaining firm characteristics (Net sales (NSALES), Operating profit (OP), Capital expenditure (CAPEX)) contained in the Scoreboard were complemented with Market capitalisation (MCAP) sourced from both the Financial Times London Share Service and Reuters. In order to create a relative measure of R&D expenditure that takes into account firm commercial size, we create the treatment variable (R&D intensity) as the ratio of the nominal R&D expenditure to net sales.

There are several categorical dummy variables indicating level of technological sophistication (low-tech, medium-low tech, medium-high tech and high-tech) that are further sub-divided into industrial sectors according to the ICB classification as well as dummy variables indicating countries. Further details on the definitions of the explanatory variables are provided in the online appendix.

The set of covariates used in our analysis is selected based on previous studies (e.g. see Hall et al., 2008), subject to their availability in our data set. In order to provide an impression on the magnitude of the main firm characteristics and their relationship to the variables of our main interest we report median values of these characteristics evaluated at each level of technological sophistication, see Table 1.

The first observation is that the number of firms belonging either to high- or medium-high-tech sectors (1367) is much larger than the number of firms belonging either to low- or low-tech sectors (292). Such an over-representation of the high-tech firms in the sample naturally reflects the original intention of collecting and maintaining the database on the world top R&D performers. In terms of employment, a median firm-specific employment is inversely proportional to the level of technological sophistication: in the high-tech sector the median employment is 4200 whereas in the low-tech sector it comprises 20,960 employees. In nominal

[5] Note, however, that data reported by the Scoreboard companies do not inform about the actual geographic distribution of the number of employees. A detailed geographic analysis should take into account the location of subsidiaries of the parent Scoreboard companies as well as the location of other production activities involved in the value-chains.

Table 1 Median values for variables by tech sector

Tech sector	*Obs.*	*EMPL (number)*	*R&D (€mln.)*	*R&D sectoral share*[a]	*R&D intensity*	*OP (€mln.)*	*NSALES (€mln.)*	*CAPEX (€mln.)*	*MCAP (€mln.)*
High	649	4200	79.8	0.532	0.115	87.6	907	40.2	2622
Medium-high	718	10,898	63.8	0.363	0.034	184.2	2368	95.6	2724
Medium-low	105	14,073	76.1	0.042	0.016	348.3	4636	206.6	5562
Low	187	20,960	74.0	0.063	0.011	567.4	9589	435.2	5929

[a]Indicates share of each tech sector in total volume of R&D expenditure

terms, the median level of R&D expenditure is about the same across the different tech sectors with a typical value of about 60–80 mln. Euro. However, the sector-specific share of R&D expenditure is not equally distributed as indicated in the column "R&D sectoral share." The lion's share of the total R&D expenditure (about 90%) is accounted for the firms in high- or medium-high-tech sectors.

As far as the treatment variable (R&D intensity) concerns, the median level is highest for the firms in the high-tech sector and it continuously decreases with the level of the technological sophistication. A median firm in the high-tech sector spends about 11.5% of its net sales volume on R&D, whereas the corresponding share for a median firm in the low-tech sector is about 1%.

It is also interesting to observe that the median values of the financial variables like operating profit, net sales, capital expenditure and market capitalisation are highest for the low-tech firms and the lowest for the high-tech firms.

5 Results

This section is sub-divided into two parts. In the first part, we report estimation results from a naive OLS regression of employment on the R&D intensity. Despite the associated econometric issues, this naive model can serve as a useful benchmark against which we can compare the results of more sophisticated methodology based on the generalised propensity score approach applied to the estimation of the functional relationship between the variables of interest, reported in the second part of this section.

OLS Estimation

The scatterplot of employment against R&D intensity is shown in Fig. 1 along with the fitted regression line. The OLS coefficient estimates are shown in the figure as well. The OLS estimate of the employment elasticity with respect to the R&D intensity is reported -0.739 indicating that a 1% increase in the R&D intensity is associated with 0.74% decrease in the number of employees. With the estimated standard error of the slope coefficient 0.026 this elasticity estimate is statistically significantly different from zero and the regression is characterised by a rather goodness of fit with the associated $R^2 = 0.327$.

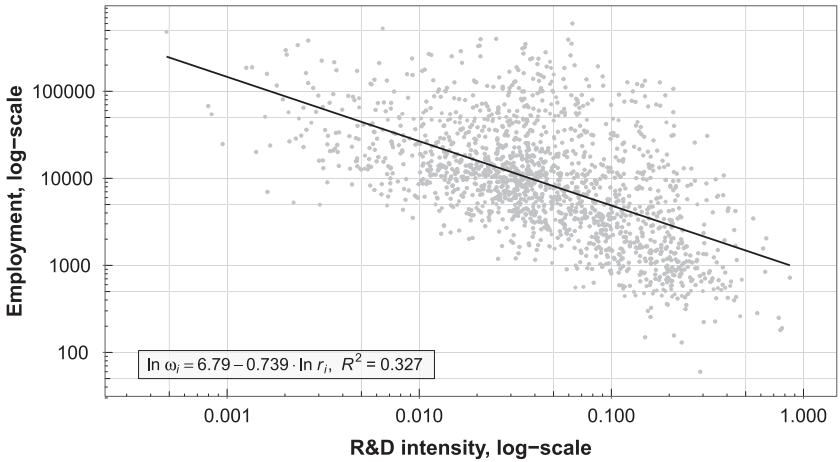

Fig. 1 OLS regression: all firms

GPS Estimation

As explained in Sect. 3 above, the application of the GPS methodology in order to estimate the dose-response function involves three steps. The results of the first step GPS estimation procedure (see Eq. (1)) are reported in Table 2. They suggest that the variation in the R&D intensity is best captured by variables such as the total capital expenditure and its square, market capitalisation and its square, as well as operating profits. Also the included industry- and region-specific dummy variables contribute substantially to the explanatory power of the first step of the GPS regression.[6] Indeed, the goodness-of-fit of this regression is quite high, yielding a R^2 of 68.2%, which is necessary in order to create a mighty propensity score able to remove biases when estimating the dose-response function between the variables of interest.

The assumption of normally distributed OLS residuals in Eq. (1) is verified by means of the Shapiro-Wilk normality test, yielding the p-value of 1.746×10^{-15}. Hence our data do not support the normality assumption. Therefore it is instructive to take a closer look at the histogram of the regression residuals, shown in Fig. 2. The fitted normal probability density function is shown as the dashed line. As seen, the residuals are characterised by too large excess kurtosis and appear to be left-skewed to be compatible with

[6] These are not shown in the regression output table in order to save the space.

Table 2 Dose regression

	Dependent variable: ln *R&D intensity*
$\ln \mathrm{CAPEX}$	-0.326*** (0.052)
$[\ln \mathrm{CAPEX}]^2$	0.185*** (0.057)
$\ln \mathrm{MCAP}$	-0.292** (0.120)
$[\ln \mathrm{MCAP}]^2$	0.315*** (0.078)
$\ln \mathrm{OP}^+$	-0.106* (0.064)
$[\ln \mathrm{OP}^+]^2$	-0.100 (0.072)
$\ln \mathrm{OP}^-$	0.030 (0.084)
$[\ln \mathrm{OP}^-]^2$	0.006 (0.013)
Constant	-1.070* (0.600)
Observations	1659
R^2	0.682

Notes: The sign-preserving log transformation of the operating profit variable was carried out as follows: for positive values $\ln \mathrm{OP}^+ = \text{if } OP > 0 : \ln OP$ and zero otherwise; for negative values $\ln \mathrm{OP}^- = \text{if } OP < 0 : -\ln(-\mathrm{OP})$ and zero otherwise. Sectoral and country dummies (not shown) were included in the regression

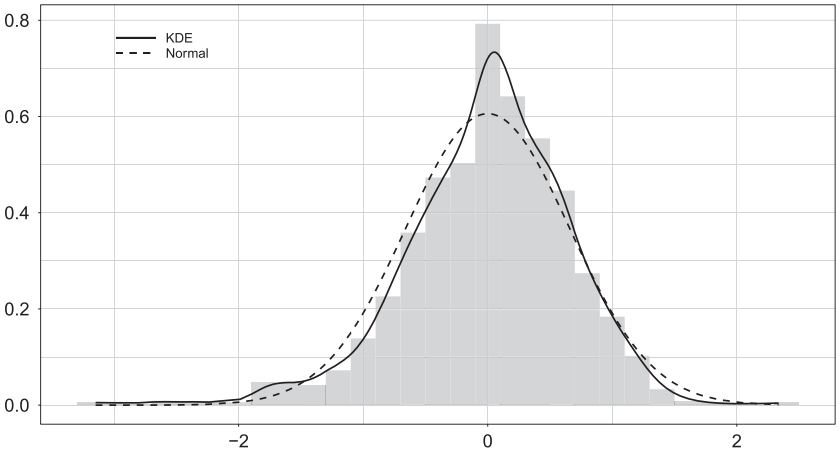

Fig. 2 Equation (1): Residuals histogram; parametric (Normal) and non-parametric kernel density estimation (KDE)

the normal distribution. Hence, instead of relying on the unfulfilled normality assumption, we estimate the GPS by means of non-parametric approach using a kernel density estimation (KDE) of the probability density function, since we have a rather large data set of 1659 observations.

The estimated non-parametric GPS is shown as the solid line in the figure. Due to its inherent flexibility, the kernel-estimated GPS matches the empirical distribution of residuals much better than the one based on the normal distribution. The GPS range is quite large [0.00289, 0.734], signifying substantial differences in the estimated propensity of the treatment level assignment across firms in our sample.

According to the estimation procedure outlined in Hirano and Imbens (2004), the next step is verification of the so-called balancing property of the GPS, preceded by imposing the common- support restriction on the data in question. The latter procedure aims to construct a more homogenised sample by filtering out aberrant observations for which propensity-score-based matching turns out problematic. The former procedure aims at testing whether conditional on observed values of the GPS variable there are no systematic differences in firms' characteristics irrespective of the assigned treatment intensity. As discussed in the online Appendix, the imposition of the common-support restriction reduced the number of firms available for the further analysis from 1659 to 1296. At the same time, the balancing property of the constructed GPS in Eq. (1) is supported by the data, see the Appendix for further details.

Next, we proceed to the estimation of the dose-response relationship between the firm innovation and employment variables. The estimation results for the second-step regression corresponding to Eq. (3) are reported in Table 3. Second step regression results clearly show that the employment response to the firm innovation (proxied by R&D expenditures) is highly non-linear, as all included polynomial terms of the latter variable report highly significant coefficients. It is also worthwhile noticing that the GPS variable enters as a statistically significant covariate both

Table 3 Conditional regression

	Dependent variable: ln *EMPL*
ln R & D intensity	−3.150*** (0.616)
$[\text{ln R \& D intensity}]^2$	−0.869*** (0.210)
$[\text{ln R \& D intensity}]^3$	−0.083*** (0.022)
GPS	−3.040*** (0.629)
GPS * ln R&D intensity	−0.885*** (0.205)
Constant	5.510*** (0.570)
Observations	1296
R^2	0.226

in levels and via the interaction term with the (log) of our treatment variable, confirming its relevance in eliminating the sample selection bias.[7] The resulting R^2 is 22.6%, which is of a comparable magnitude reported in other studies (Egger et al., 2012).

In order to facilitate the interpretation of the estimation results, we have plotted the estimated dose-response and marginal treatment effect functions in the upper and middle panels of Fig. 3, respectively. The bands around the estimated functions are 95% bootstrap confidence intervals. Observe that in order to facilitate the description of the results in the lower panel of the figure we have plotted the cumulative share of employment in the firms in our sample as a function of the R&D intensity. The curve in the lower panel reveals that 90% of employment in our data sample is accounted by firms with the R&D intensity in the interval between 0.6% and 15%. There are 1088 out of 1296 firms, or about 84%, of the total sample in this interval. There are 21 and 187 firms in the left and right 5% tails of the cumulative employment distribution sorted by the R&D intensity.

The shape of the estimated dose-response function is generally downward sloping, which is broadly consistent with the naive OLS estimation results reported in Sect. 5. However, recall that according to the OLS results the estimated employment elasticity is uniformly negative at all R&D intensity levels. In contrast, the estimated dose-response function using the GPS suggests that the magnitude of the response of employment to changes in the R&D intensity varies with the level of the firms' innovation intensity. This non-linearity in the employment response is well illustrated by the marginal treatment effect function, which can be interpreted as employment elasticity with respect to R&D intensity, that is shown in the middle panel of Fig. 3.

The estimated elasticity of interest has a hump-shaped form. Hence it is convenient to summarise our findings by distinguishing between different treatment intensity levels taking into the consideration such hallmarks as the top and bottom 5% cumulative employment thresholds. For relatively low treatment intensity levels (below 0.6%) the employment elasticity increases in the absolute value from −0.5% up to about −1.5% as the treatment intensity falls. However, given a rather small number of observations

[7] Higher order power transformations of the GPS variable turned out to be insignificant and therefore were omitted from the model specification for the sake of parsimony.

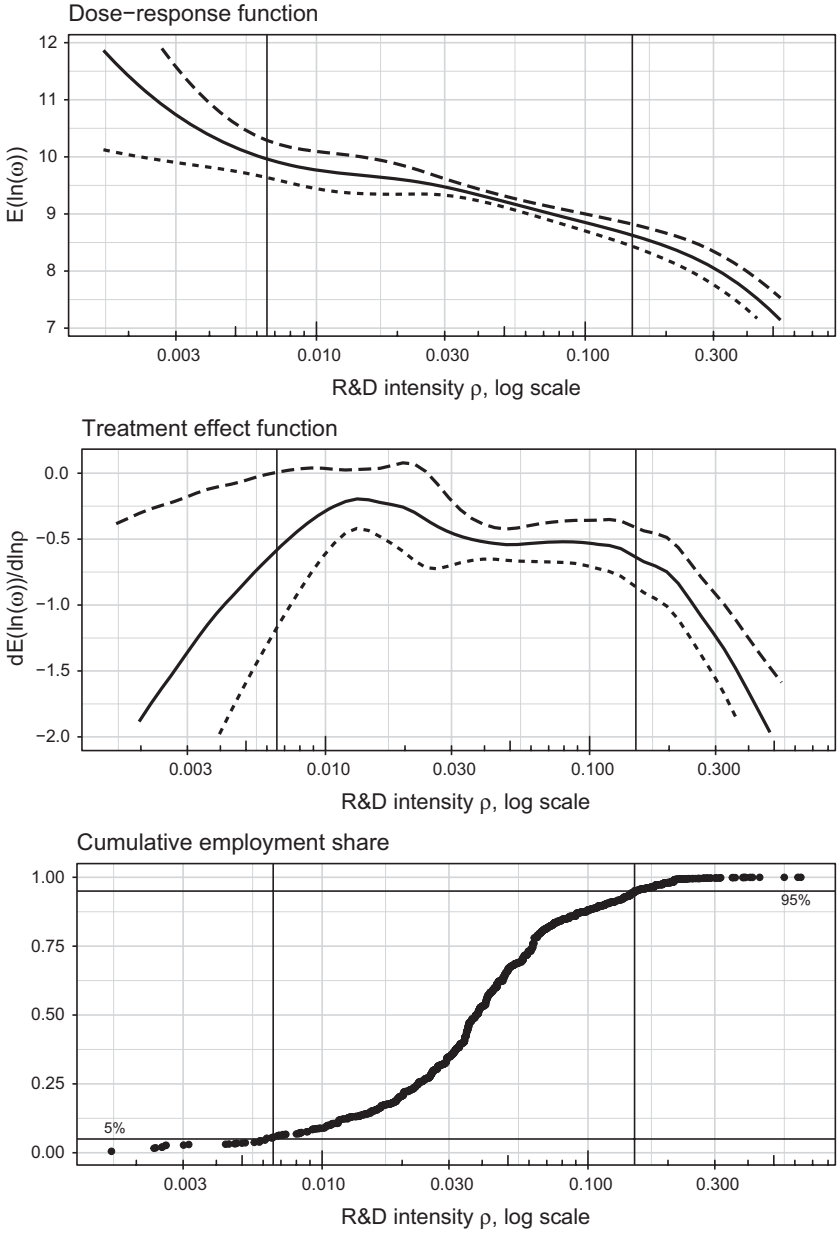

Fig. 3 Dose-response, treatment effect functions and cumulative employment share

in this part of the distribution these estimate values have to be taken with caution.

For the firms within the central 90% interval of the treatment intensity one can make the following two observations. First, for the firms with R&D intensity in the interval between 0.6% and 3% the estimated elasticity is not significantly different from zero, as the bootstrapped 95% confidence interval includes zero line. This suggests a labour-neutral effect of innovation for the firms with medium-low and medium levels of innovation intensity. Second, for the firms with the medium-high levels of the R&D intensity pertaining to the interval between 3% and 15% the estimated elasticity is negative and significantly different from zero. For these firms it is estimated around −0.5% with the associated 95% confidence interval about (−0.3%, −0.7%), suggesting labour-saving effect of innovation. Notwithstanding that this value is substantially lower than that reported by the OLS estimation (−0.74%) earlier in the text.

Turning to the firms with the highest R&D intensity (>15%), this labour-saving effect turns out to be even more pronounced. In this interval, the estimated employment elasticity gradually increases (in the absolute magnitude) from −0.5% to −2.0%, suggesting that the innovation leaders tend to react more and more disproportionately stronger to changes in the R&D intensity in reducing their labour force than innovation followers and moderate innovators.

All in all, our estimation results when compared to those from the naive OLS regression suggest that the employment effect of innovation varies with the level of technological sophistication and warrant against application of estimation techniques that does not accommodate such level dependence. For the firms with rather low to medium ratios of R&D expenditure to net sales this effect tends to be overestimated by the OLS regression whereas understated for the firms on the other side of the spectrum characterised by high values of R&D intensity.

It is instructive to compare our results with traditional point estimates available in the previous literature, despite the fact that studies summarised in Sect. 2 focus on the employment elasticity with respect to a nominal measure of the R&D expenditure, whereas we focus on the employment elasticity with respect to a relative measure of the R&D expenditure. Our results, emphasising the complexity of the non-linear relationship between employment and innovation, are complementing those of Bogliacino (2014), who equally finds that R&D investment expenditures have a non-linear effect on the firm employment, depending on the R&D intensity.

However, compared to the most of the published literature, our results reveal no support for a job-creating aspect of innovation at least when the world top R&D performers are scrutinised. For this particular sub-sample of firms we find that the effect of innovation is at best labour-neutral at the relatively low values of the R&D intensity. For higher innovation intensity levels, the labour-saving effect of innovation becomes increasingly pronounced, as knowledge-intensive firms are looking for high-skilled labour force which is typically in much shorter supply and correspondingly more expensive than their low-skilled fellows.

6 Conclusions, Policy Recommendations and Limitations

The objective of the study is to expose the entire innovation-employment relationship for different R&D intensity levels in a continuous framework. We use a large international firm-level panel data set for OECD countries and employing a flexible semi-parametric method—the generalised propensity score—allows us to estimate the full functional relationship between the R&D-driven innovation and firm employment as well as address important econometric issues, which is not possible in the standard estimation approach used in the previous literature. This is our main contribution to the academic literature and policy debate; to the best of our knowledge no comparable studies analysing the employment effect of innovation in a continuous setting are available in the literature.

In order to answer these questions, we have based our empirical microeconometric analysis on a large international firm-level panel dataset for OECD countries, and our proxy for technology has been a measurable and continuous variable, while the majority of previous studies have relied on either indirect proxies of the technological change or dummy variables (such as the occurrence of the product and process innovation). In particular, we have employed the EU Industrial R&D Investment Scoreboard data set for 2500 R&D performers worldwide. In addition to firm-level innovation expenditures, we have used also economic and financial variables, which allowed us to control for important firm-specific effects, along with sectoral and regional dummies.

Our results suggest that a care should be taken when analysing employment-innovation nexus. Depending on the level of R&D intensity, we find that the innovation impact on employment can be negative

too—findings that have been reported also in previous studies. This labour-saving aspect of innovation is more pronounced for firms with medium-high levels of R&D intensity and it tends to increase with the levels of R&D intensity. In terms of policy recommendation, our results imply that these companies should not be immediately targeted by policies aiming to achieve both innovation and employment targets of the R&I Strategy in the same time.

Turning to limitations of our study, an important caveat of our empirical analysis concerns the nature of the Scoreboard sample. First, while other data sets, such as the OECD BERD data, can be considered as fully representative of OECD economies, in the EU Industrial R&D Investment Scoreboard data used in this study only R&D "champions" are considered. This is a clear limitation of our data, the results of which cannot be straightforwardly extrapolated to, for example, SMEs.

A further limitation of the data used in our study is that R&D Scoreboard data do not allow us to identify the effect of product and process innovations separately. However, as discussed in the introduction, the employment effect of innovation can be very different depending on the nature of innovation. In order to separately identify the employment effect of the product and process innovation, other sources of data, such as the Community Innovation Survey (CIS), need to be used, which is a promising area for the future research.

Lastly, in our study we focus on the snapshot of the economy at one period of time without taking higher order effects of firms innovation activity. In the longer run, investing in the innovation activity encourages knowledge-based economy, drives demand for high-skilled, educated workers and eventually brings a country on the higher growth path. However, a comprehensive assessment of these effects is only possible within general-equilibrium models that capture vertical and horizontal linkages between firms, which is not possible to account for in micro-econometric studies, such as the one presented in this paper (Kancs & Ciaian, 2011; Brandsma & Kancs, 2016). Hence aligning our results with macro results is indeed important for enhancing our understanding of the employment effect of innovation and it sets a promising avenue for the future research.

Acknowledgments The authors acknowledge helpful comments from Francesco Bogliacino and Daria Ciriaci, Mihails Hazans, Enkelejda Havari, Thomas Hemmelgarn, Giuseppe Piroli, the editor of the journal, two anonymous referees,

participants of the conferences on Counterfactual Methods for Policy Impact Evaluation in Rome and Milan, the Scottish Economic Society Annual Conference in Perth and the 57th Meeting of the Italian Economic Association in Milan as well as participants of the research seminars at the European Commission, Bank of Latvia, University of Leuven and Stockholm School of Economics in Riga. We are grateful to Pietro Moncada-Paterno-Castello and the IRI team for granting access to the EU Industrial R&D Investment Scoreboard data. This article follows the methodological framework of Kancs and Siliverstovs (2019). The authors are solely responsible for the content of the paper. The views expressed are purely those of the authors and may not under any circumstances be regarded as stating an official position of the European Commission or the Bank of Latvia. Computations were performed in ***R*** (R Development Core Team, 2008).

Appendix

Explanatory Variables

The following groups of the explanatory variables were used in our analysis in the main text:

- *Net sales*, NSALES: In line with the accounting definition of sales, sales taxes and shares of sales of joint ventures & associates are excluded. For banks, sales are defined as the "Total (operating) income" plus any insurance income. For insurance companies, sales are defined as "Gross premiums written" plus any banking income.
- *Operating profit*, OP: Profit (or loss) before taxation, plus net interest cost (or minus net interest income) and government grants, less gains (or plus losses) arising from the sale/disposal of businesses or fixed assets. Due to the fact that companies report both positive and negative operating profit, we cannot take a logarithmic transformation of this variable. In order to do so, we created the following two variables $\ln \mathrm{OP}^{+}$ and $\ln \mathrm{OP}^{-}$. The former variable is equal to the log of actual values whenever a firm reports positive profit and zero otherwise. The latter variable is equal to the log of absolute actual values multiplied by minus one whenever a firm reports negative profit and zero otherwise.
- *Capital expenditure*, CAPEX: The expenditure used by a company to acquire or upgrade physical assets such as equipment, property, industrial buildings. In company accounts capital expenditure is added to the asset account (i.e. capitalised), thus increasing the

amount of assets. It is disclosed in accounts as additions to tangible fixed assets.

- *Market capitalisation*, MCAP: The share price multiplied by the number of shares issued at a given date. Market capitalisation data have been extracted from both the Financial Times London Share Service and Reuters. These reflect the market capitalisation of each company. The gross market capitalisation amount is used to take into account those companies for which not all the equity is available on the market.
- *Country dummies*: There are 36 distinct countries included in the estimation sample.
- *Industry sector dummies*: The industry sectors are based on the ICB classification. The level of disaggregation is generally the three-digit level of the ICB classification, which is then converted to NACE Rev.2.
- *Sectoral dummies*: In order to account for the sectoral heterogeneity with respect to the R&D intensity, we regroup all firms into four sub-samples according to the level of their technological sophistication. Following the OECD classification, all firms in our sample are regrouped into four 3-digit Industry Classification Benchmark (ICB) groups: high-, medium-high-, medium-low- and low-tech companies:
  - *High-tech*: Technology hardware & equipment, Software & computer services, Pharmaceuticals & biotechnology, Health care equipment & services, and Leisure goods;
  - *Medium-high-tech*: Industrial engineering, Electronic & electrical equipment, General industrials, Automobiles & parts, Personal goods, Other financials, Chemicals, Aerospace & defence, Travel & leisure, Support services, and Household goods & home construction;
  - *Medium-low-tech*: Food producers, Fixed line telecommunications, Beverages, General retailers, Alternative energy, Media, Oil equipment, services & distribution, and Tobacco;
  - *Low-tech*: Gas, water & multi-utilities, Oil & gas producers, Nonlife insurance, Industrial metals & mining, Construction & materials, Food & drug retailers, Banks, Electricity, Industrial transportation, Mobile telecommunications, Forestry & paper, Mining, Life insurance.

Verification of GPS Balancing Property

The balancing property of the constructed GPS variable is verified following the procedure suggested by Hirano and Imbens (2004). Each covariate is subdivided into three groups of 553 observations according to the percentiles of the distribution of the treatment intensity variable. The initial testing of the balancing property amounts to testing whether the average value of a particular variable in every group is equal to the average value in the remaining groups. The results of these tests are reported in Table 4. Only for a handful of covariates we cannot reject the tested null hypothesis at usual significance levels, indicating that there is a strong heterogeneity among covariates belonging to these three groups pertinent to different values of the treatment intensity. A well-specified GPS should be able to successfully account for these differences.

Before verifying the balancing properties of the GPS, we impose the so-called common-support restriction. The purpose of this restriction is to filter out observations that are rather dissimilar in their characteristics when used for the GPS computation in the first step, see Eqs. (1) and (2). As argued by Becker et al. (2012) , it is advisable to impose the common-support condition in order to improve the balancing properties of the GPS and hence achieve more reliable estimation results.

For each treatment group, defined above, $k = 1, 2, 3$, we evaluate GPS values for each observation i at the respective median treatment value, $GPS_{i_j}^{k}$. We determine the common support region by comparing values of GPS_i^j for each $j = k$ with those computed for other groups $j \neq k$ at the median treatment value of the selected group j. Those observations for which $GPS_i^{j \neq k}$ that fall outside of the range of GPS_i^j are labelled as those

Table 4 Covariate balance, t-statistics (initial data)

	Dose group 1	*Dose group 2*	*Dose group 3*
$\ln \mathrm{CAPEX}$	-18.35	-1.86	19.32
$[\ln \mathrm{CAPEX}]^2$	-14.91	-0.28	17.44
$\ln \mathrm{MCAP}$	-8.93	2.01	6.19
$[\ln \mathrm{MCAP}]^2$	-8.11	2.28	5.44
$\ln \mathrm{OP}^+$	-14.36	-3.91	15.80
$[\ln \mathrm{OP}^+]^2$	-13.13	-1.22	14.03
$\ln \mathrm{OP}^-$	-5.08	-6.13	8.76
$[\ln \mathrm{OP}^-]^2$	-3.17	-4.62	6.69

Table 5 Number of observations in common support

Common support	*Dose group 1*	*Dose group 2*	*Dose group 3*
1296	1651	1430	1411

Table 6 Number of observations by dose group and block

	1:(Total)	*1*	*¯1*	*2:(Total)*	*2*	*¯2*	*3:(Total)*	*3*	*¯3*
m(GPS) = 1	738	68	670	492	105	387	776	87	689
m(GPS) = 2	200	68	132	230	104	126	212	87	125
m(GPS) = 3	146	68	78	207	104	103	97	86	11
m(GPS) = 4	112	68	44	184	104	80	110	87	23
m(GPS) = 5	100	68	32	183	105	78	101	87	14

that do not satisfy the common support restriction and therefore are removed from the analysis. At the final step, we retain only observations i that survive the common support filtering in all treatment groups. In Table 5 we report the number of observations retained in each group that satisfy the common support condition. Taken together, only 1296 out of 1659 observations can be considered as comparable in terms of their characteristics and hence are retained for a further analysis.

In order to check whether the balancing property of the constructed GPS can be warranted in our data, we subdivide each group into blocks of approximately the same size corresponding to quintiles of the respective GPS. The resulting cell sizes of each block are reported in Table 6. The testing procedure of the differences in means for each variables and for each treatment group conditioning on the GPS values is conducted in the following two steps. In the first step, five tests for the differences in means are conducted for each block. Then in the second step the computed block-specific differences in means are combined using the total number of observations in each block as weights. The balancing properties of covariates adjusted for the GPS are reported in Table 7. Compared to the results for unadjusted covariates reported in Table 4, a substantial improvement can be observed, as only three test statistics exceed the nominal 5% significance level. Hence, we can conclude that the generalised propensity score is appropriately defined.

Table 7 Covariance balance, t-statistics (GPS-adjusted)

	Dose group 1	*Dose group 2*	*Dose group 3*
$\ln \text{CAPEX}$	-2.43	-0.51	2.12
$[\ln \text{CAPEX}]^2$	-1.52	-0.58	1.84
$\ln \text{MCAP}$	-0.84	0.85	0.33
$[\ln \text{MCAP}]^2$	-0.70	0.82	0.18
$\ln \text{OP}^+$	-0.92	-0.47	1.50
$[\ln \text{OP}^+]^2$	-0.98	-0.28	1.45
$\ln \text{OP}^-$	0.36	-0.87	0.06
$[\ln \text{OP}^-]^2$	0.35	-0.84	0.16

References

Antonucci, T., & Pianta, M. (2002). Employment effects of product and process innovation in Europe. *International Review of Applied Economics, 16*(3), 295–307.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*(2), 277–297.

Becker, S. O., Egger, P. H., & von Ehrlich, M. (2012). Too much of a good thing? On the growth effects of the EU's regional policy. *European Economic Review, 56*(4), 648–668.

Bia, M., Flores, C. A., & Mattei, A. (2011). *Nonparametric estimators of dose-response functions.* LISER Working Paper Series 2011-40. Luxembourg Institute of Socio-Economic Research.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*(1), 115–143.

Bogliacino, F. (2014). Innovation and employment: A firm level analysis with European R&D Scoreboard data. *Economia, 15*(2), 141–154.

Bogliacino, F., Piva, M., & Vivarelli, M. (2012). R&D and employment: An application of the LSDVC estimator using European microdata. *Economics Letters, 116*(1), 56–59.

Bogliacino, F., & Vivarelli, M. (2012). The job creation effect of R&D expenditures. *Australian Economic Papers, 51*(2), 96–113.

Brandsma, A., & Kancs, D. (2016). RHOMOLO: A dynamic general equilibrium modelling approach to the evaluation of the EU's R&D policies. *Regional Studies, 49*, 1340–1359.

Bruno, G. S. F. (2005). Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals. *Stata Journal, 5*(4), 473–500.

Bun, M. J. G., & Kiviet, J. F. (2003). On the diminishing returns of higher-order terms in asymptotic expansions of bias. *Economics Letters, 79*(2), 145–152.

Ciriaci, D., Moncada-Paterno-Castello, P., & Voigt, P. (2013). Innovation and job creation: A sustainable relation? Working Papers on Corporate R&D and Innovation 2013-01, European Commission, DG Joint Research Centre.

Egger, P. H., von Ehrlich, M., & Nelson, D. R. (2012). Migration and trade. *The World Economy, 35*(2), 216–241.

Hall, B. H., Lotti, F., & Mairesse, J. (2008). Employment, innovation, and productivity: evidence from Italian microdata. *Industrial and Corporate Change, 17*(4), 813–839.

Heinrich, C., Maffioli, A., & Vázquez, G. (2010). Impact-evaluation guideline. Technical Notes No. IDB-TN-161, Inter-American Development Bank.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. In A. Gelman & X.-L. Meng (Eds.), *Applied Bayesian modeling and causal inference from incomplete-data perspectives* (pp. 73–84). Chichester: Wiley.

Ivanova, O., Kancs, D., & Thissen, M. (2019). EU Economic Modelling System: Assessment of the European Institute of Innovation and Technology (EIT) Investments in Innovation and Human Capital, Publications Office of the European Union, Luxembourg. https://doi.org/10.2791/184008

Kancs, D., & Ciaian, P. (2011). Modelling the flow of knowledge and human capital: A framework of innovative capital. *International Journal of Public Policy, 7*(1), 134–160.

Kancs, D., & Siliverstovs, B. (2016). R&D and non-linear productivity growth. *Research Policy, 45*, 634–646.

Kancs, D., & Siliverstovs, B. (2019). Employment Effect of Innovation. JRC Working Papers in Economics and Finance 2019/2, Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/52757. JRC114588

Lachenmaier, S., & Rottmann, H. (2007). Employment effects of innovation at the firm level. *Journal of Economics and Statistics, 227*(3), 254–272.

Marx, K. (1867). *Das Kapital, Kritik der politischen Oekonomie*. Hamburg: Erster Band.

Moncada-Paterno-Castello, P., Ciupagea, C., Smith, K., Tubke, A., & Tubbs, M. (2010). Does Europe perform too little corporate R&D? A comparison of EU and non-EU corporate R&D performance. *Research Policy, 39*(4), 523–536.

Pianta, M. (2004). Innovation and employment. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *Handbook of innovation* (Chapter 22, pp. 568–598). Oxford University Press.

R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–50.

Smolny, W. (1998). Innovations, prices and employment: A theoretical model. *Journal of Industrial Economics, 46*(3), 359–381.
Van Reenen, J. (1997). Employment and Technological Innovation: Evidence from U.K. Manufacturing Firms. *Journal of Labor Economics, 15*(2), 255–84.
Young, A. (1993). Substitution and complementarity in endogenous innovation. *Quarterly Journal of Economics, 108*(3), 775–807.

Structural Transformation, Biased Technological Change and Employment in Vietnam

Philip Abbott, Finn Tarp, and Ce Wu

1 Introduction

Between 2000 and 2011 GDP grew on average by 6.9 per cent per year in Vietnam, while employment increased at only 2.7 per cent per year. Two perspectives can be taken on these outcomes. On the one hand, employment has been growing faster than population, and even faster GDP

European Journal of Development Research (2017) 29, 54–72. https://doi.org/10.1057/ejdr.2015.64; published online 29 October 2015.
The online version of this chapter is available Open Access.

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_16

growth allows for increasing per capita income. McCaig and Pavcnik (2013) argue that this is a case of successful development. On the other hand, prominent Vietnamese policymakers are concerned that employment growth has been too modest, and that slow employment generation, particularly in the modern sectors, causes slower wage growth, creation of too few high-quality jobs and serious underemployment problems, especially in rural areas. A study by Vietnam's labour ministry done with the ILO argues this case, and asserts that problems of employment generation are likely to increase in the future (MOLISA and ILO, 2010). The authors of that report largely attribute 'slow employment growth' to inadequate structural transformation. They repeat the common though controversial argument that the Asian miracle as implemented in Vietnam is based on heavy capital investment with no technical change (for example, Kim & Lau, 1994; Krugman, 1994; Young, 1995; Collins & Bosworth, 1996). They raise the concern that development in Vietnam may have been excessively capital intensive, a characteristic particularly pronounced in investments made by state-owned and foreign-invested enterprises. High and growing minimum wages may in this line of thinking have fostered capital-intensive growth, working against Vietnam's comparative advantage as a labour-abundant, low wage economy. The MOLISA/ILO (2010) study concludes that the investment structure and labour policies must be adjusted so that high value-added industries, and particularly private firms, can be better positioned to create more jobs.

We rely on data for seven aggregate sectors and the overall Vietnamese economy from 2000 to 2011 provided by GSO (2014) to examine the differences observed between employment and GDP growth. A decomposition framework is used to explore three hypotheses that might explain employment growth in Vietnam: (i) structural transformation; (ii)

P. Abbott (Deceased)
Purdue University, West Lafayette, IN, USA

F. Tarp (✉)
UNU-WIDER, Helsinki, Finland

University of Copenhagen, Copenhagen, Denmark
e-mail: finn.tarp@econ.ku.dk

C. Wu
Indiana Wesleyan University, Marion, IN, USA

technological progress and (iii) excessively capital-intensive development. Two institutional biases that facilitate capital-intensive development are studied: (a) minimum wage policy that may distort the wage-rental ratio and (b) state as well as foreign investment that may be overly capital intensive. Our results show that only about one-third of the difference in employment and GDP growth is because of shifts from low- to higher-productivity sectors. The remaining two-thirds come from declining labour use per unit output that is also found in agriculture. Labour-augmenting technical change[1] is an important factor behind Vietnam's successful growth, and restructuring from capital-intensive state-owned enterprises is best seen as a missed opportunity, given very small changes in the state's employment share.

Section 2 explores further the context shaping ongoing debates on labour policy in Vietnam. Section 3 reports our decomposition of employment growth, while Section 4 considers the impacts of biased technological change on employment. Section 5 examines institutional biases—minimum wage policy and investment biases—that might influence employment growth. Section 6 concludes.

2 Context and Background

Three issues are examined here that relate to Vietnam's recent structural transformation and employment growth. First, Vietnam's recent economic history is explored. Next, its performance is compared with outcomes in other Asian economies. How this experience relates to literature on what is behind the Asian miracle is then considered.

Structural Transformation

For the past two decades Vietnam has experienced rapid GDP growth. Table 1 presents the sectoral composition of GDP and employment in Vietnam for 2007, as well as sectoral growth rates from 2000 to 2007 and 2007 to 2011 (GSO, 2014).[2] This division was chosen to reflect slowdowns

[1] We will use the terminology 'labour-augmenting technical change', corresponding with the notion that efficiency of the labour input is increasing. 'Labour-saving' is ambiguous, since an improvement in capital efficiency could reduce labour demand, but only if the elasticity of substitution exceeds one.

[2] Though no distinction is indicated on the GSO (2014) website, data for employment clearly include informal labour given the large reported size of the agricultural sector, and the

Table 1 GDP and employment in Vietnam

	2007		*2000–2007*	*2007–2011*
	Trillion Dong	*Share—%*	*Growth—%/year*	
GDP	461	–	7.5	5.9
Agriculture	83	18	3.7	3.3
Manufacturing	113	25	11.2	7.0
Construction	43	9	10.4	4.7
Infrastructure services	106	23	7.1	7.4
Education & health	22	5	7.4	6.9
Energy & natural resources	38	8	6.2	4.2
Other services	57	12	6.6	6.0
	Million persons			
Employment	45.2	–	2.6	2.7
Agriculture	23.9	53	-0.3	0.4
Manufacturing	5.7	13	6.7	5.2
Construction	2.4	5	11.8	7.7
Infrastructure services	7.4	16	3.5	5.7
Education & health	1.9	4	6.3	3.8
Energy & natural resources	0.4	1	3.1	-0.1
Other services	3.6	8	15.2	2.5

Source: GSO (2014)

in economic growth in Vietnam and elsewhere following the global recession (Abbott & Tarp, 2012). State intervention in reaction to the great recession may have influenced employment trends. Sectoral employment shares and sectoral employment growth rates are also reported. The table clearly reflects a case of structural transformation at work. From 2000 to 2007 GDP grew at 7.5 per cent per year. Agriculture grew at 3.7 per cent, while manufacturing grew at 11.2 per cent, construction at 10.4 per cent and various services at over 6.6 per cent per year. The slower GDP growth from 2007 to 2011, at 5.9 per cent per year, brought greater reductions in manufacturing (to 7.0 per cent) and construction growth (to 4.7 per cent) than for agriculture (to 3.3 per cent). From 2000 to 2007, employment in agriculture actually declined, while manufacturing employment increased at 6.7 per cent and construction at 11.8 per cent. While slower GDP growth after 2007 brought slower employment growth in

numbers reported for manufacturing, construction and infrastructure services relative to numbers from the Enterprise survey (GSO, 2011) that looks only at formal firms. When we look at wage data from that survey later, that information is for formal firms, and will be so noted.

manufacturing and especially construction, employment growth in lower-productivity agriculture and some service sectors increased, and hence overall employment growth held roughly constant. Thus, problematic structural transformation appears to be a smaller issue after 2007 than before. Differences in value added versus employment growth rates are noticeable for agriculture and manufacturing and in some services but not elsewhere. Infrastructure services (transportation, communications, hotels and tourism, and wholesale and retail trade) exhibit slower employment growth, whereas other services (professional and public services) have shown rapid employment growth, especially after 2007.

Figure 1 compares labour productivity across sectors in Vietnam. It shows the typically large differences in productivity levels, with manufacturing and construction yielding the highest productivity rates, while productivity in agriculture is much lower. In addition to showing sectoral differences, Figure 1 also demonstrates that productivity has been

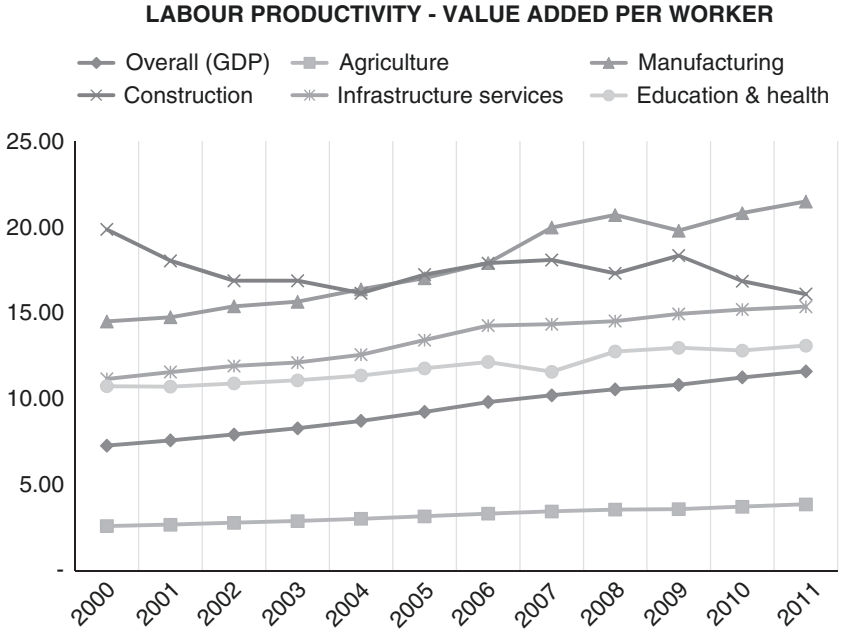

Fig. 1 Sectoral labour productivity in Vietnam. (Notes: Real value added per worker is measured in million dong (at 1994 prices) per year. Source: GSO (2014))

improving steadily in most sectors in Vietnam, averaging 4.2 per cent per year from 2000 to 2011. In contrast to a simplistic structural transformation story, agricultural productivity has been improving at 3.6 per cent per year. Manufacturing productivity also grew at 3.6 per cent per year. The faster growth for overall productivity reflects in part sectoral share changes towards higher-productivity activities, but may also capture capital-labour substitution or biased technical change.

Investment accounted for 34–42 per cent of Vietnam's GDP in 2000–2011, but investment efficiency is relatively low, suggested by high incremental capital-output ratios, especially after 2007 (GSO, 2014). In addition to lowering employment growth, capital-intensive development may have contributed to sluggish restructuring of the economy from rural to urban areas, from agriculture to manufacturing, and from public to private firms, resulting in slower than desired labour migration out of sectors where productivity is low (MOLISA and ILO, 2010).

International Comparison

While Vietnam is in some ways unique, the above trend in labour demand relative to GDP growth is found throughout East Asia. Table 2 compares several Asian countries. The elasticity of labour demand (employment) growth due to GDP growth around 2006–2008 was 0.32 in Vietnam, while it was around 0.1 in faster-growing China. Values for Thailand and South Korea are also lower than in Vietnam in recent years. Estimates very similar to those for Vietnam are found in several other East Asian countries. This measure is used by the ILO to reflect concerns similar to those voiced in Vietnam on employment generation (Kapsos, 2005), and is coupled with the concern that development may be overly capital intensive throughout Asia.

Global trends in GDP and employment growth have inspired a recent literature on 'premature deindustrialization' (Rodrik, 2015). The second step in structural transformation, following a first shift from agriculture to manufacturing, is the move from a manufacturing-based economy to greater emphasis on services, where wages are often lower. The proponents of premature deindustrialization see this happening sooner and at lower GDP per capita in East Asia and in recently developing countries than was the case for the now developed economies. While Vietnam, given its sustained manufacturing employment growth, is not at this point counted among the most problematic cases, the pessimistic perspective

Table 2 Real GDP growth and employment growth in selected Asian countries, 1986–2008

Country	*1986–1990*	*1991–1995*	*1996–2000*	*2001–2005*	*2006–2008*
Gap between real GDP growth and employment growth (%)					
China	5.7	11.0	7.5	8.6	10.5
Korea, Republic	6.4	5.6	3.4	3.4	3.5
Thailand	7.7	8.3	-0.4	3.4	3.1
Indonesia	4.6	5.5	-1.9	2.7	4.2
Philippines	2.0	-0.9	1.7	2.0	3.3
Vietnam	2.7	6.4	4.7	5.7	5.0
Elasticity of labour demand with respect to real GDP[a]					
China	0.28	0.11	0.13	0.1	0.06
Korea, Republic	0.33	0.29	0.25	0.24	0.16
Thailand	0.25	0.04	1.58	0.33	0.25
Indonesia	0.35	0.30	2.91	0.42	0.30
Philippines	0.57	1.42	0.58	0.55	0.40
Vietnam	0.53	0.28	0.30	0.27	0.32

[a]The elasticity of labour demand with respect to real GDP is calculated as the ratio between the percentage change in employment and the percentage change in real GDP, measuring the responsiveness of employment to economic growth

Source: Calculated using data from World Bank (2013)

among its proponents reflects concerns that this may be an issue in the future.

Technical Change and the Asian Miracle

The role of technological progress during the course of Vietnam's development in the past two decades tends in our assessment to be discounted. According to MOLISA and ILO (2010), total factor productivity (TFP) growth only accounted for 26 per cent of GDP growth, whereas capital accumulation contributed to more than 60 per cent of GDP growth. Previous studies such as ILSSA (2008) also found low TFP growth rates.

The role of technological progress during the course of economic growth in East Asia is controversial. Some economists (for example, Kim & Lau, 1994; Krugman, 1994; Young, 1995; Collins & Bosworth, 1996) maintain that East Asian growth is mainly because of capital accumulation and that technological progress is unimportant. Others (for example, World Bank, 1993; Pack & Page, 1994; Freeman, 1995; Rodrik, 1997; Nelson & Pack, 1999; Krüger et al., 2000) argue that technological

progress, particularly labour-saving technological change, fuelled the extraordinary economic growth that constituted the 'Asian miracle'. Moreover, Rodrik (1997) explicitly points out that the low TFP growth rates found in most studies are problematic. They fail to incorporate biased technological progress in their calculations.

3 Employment Growth Decomposition

We now turn to employment growth decomposition to quantitatively analyse the respective contributions of biased technological change, structural transformation and institutional biases to employment growth for each sector and for the overall Vietnamese economy. The decomposition starts with a straightforward relationship that links employment, labour efficiency, sectoral output (value added) shares and economic output:

$$L_t = \sum_i L_{it} = \sum_i a_{Lit} S_{it} Y_t \tag{1}$$

where L_t is total employment at time t, which is the sum of sectoral employment L_{it} across the i sectors. a_{Lit} denotes the unit labour efficiency coefficient for sector i at time t, which is the ratio between labour used in sector i and output in sector i. S_{it} is the proportion of output in sector i in total output at time t, and Y_t is the output (GDP) of the entire economy at time t. Differentiating equation (1) with respect to time gives:

$$\frac{dL_t}{dt} = \sum_i S_{it} Y_t \left(\frac{da_{Lit}}{dt}\right) + \sum_i a_{Lit} Y_t \left(\frac{dS_{it}}{dt}\right) + \sum_i a_{Lit} S_{it} \left(\frac{dY_t}{dt}\right) \tag{2}$$

Defining the growth rate of variable X as $gX = (dX/dt) \cdot (1/X)$, we turn equation (2) into growth rates as follows:

$$gL_t = \sum_i \frac{a_{Lit} S_{it}}{a_{Lt}} ga_{Lit} + \sum_i \frac{a_{Lit} S_{it}}{a_{Lt}} gS_{it} + gY_t \tag{3}$$

where $a_{Lt} = L_t/Y_t = \Sigma_i a_{Lit} S_{it}$, the overall labour-output ratio for the economy.

Equation (3) incorporates the factors that may contribute to slower than desired employment growth as discussed earlier. Biased technical

change affects employment by altering input efficiency. Minimum wage policy may distort input-output ratios by altering relative input prices, and therefore relative factor intensities. Thus, the first term on the right-hand side of equation (3) measures the combined contribution of two effects—capital-labour substitution and technical change—both of which are incorporated in the changes in the labour-output coefficient (a_{Lit}). Structural transformation fundamentally influences sectoral output shares, with traditional sectors shrinking and modern sectors expanding. State investment policies direct resources to flow into specific sectors, and hence may change sectoral output shares, Sit, as well. The employment effects of both structural transformation and state investment policies are summarized in the changes in sectoral output shares, gS_{it}. Therefore, the second term in equation (3) measures the contribution of structural transformation and state investment bias to sectoral employment. If labour-output coefficients, a_{Lit}, and output shares, S_{it}, remain constant, one would observe the same growth rates for output and employment ($gL_t = gY_t$). The first two terms in equation (3) are weighted by relative sectoral labour efficiency and sectoral output.

To implement our decomposition we assembled annual data from the GSO website (2014) over the years 2000–2011 for both employment a (L_{it}) and sectoral output. The data were originally disaggregated at an 18-sector level. Since structural transformation mainly focuses on shifts between aggregate sectors, such as agriculture, manufacturing and services, we aggregated the data into seven sectors, mainly eliminating details in the service sectors. Among the seven economic sectors, there is an aggregate traditional sector (agriculture), a manufacturing sector, an energy and natural resources sector, and four service sectors (including construction). The row titled 'GDP' represents the overall economy.

The results of our employment growth decomposition analysis are presented in Table 3. Growth rates for sectoral output shares (gSi), sectoral employment, gLi, and labour productivity (labour—output ratios), gaLi, are reported for 2000–2007 and 2007–2011. The columns labelled 'BiasTech' are the contributions to labour employment growth because of improved labour productivity, and hence the components of the first term of equation (3). These may be because of technical progress or minimum wage bias. The column labelled 'StrCh' is the contribution because of sectoral share changes, and hence the components of the second term of equation (3). These may be because of structural transformation or

Table 3 Employment growth decomposition

Sectors	*2000–2007 (%)*		*gaLi*	***Bias Tech***	*StrCh*	*gY*	*2007–2011 (%)*		*gaLi*	***Bias Tech***	*StrCh*	*gY*
	gSi	*gLi*					*gSi*	*gLi*				
Agriculture	-3.7	-0.3	-4.1	-2.39	-2.20	–	-2.6	0.4	-2.8	-1.42	-1.34	–
Manufacturing	3.8	6.7	-4.6	-0.50	0.41	–	1.1	5.2	-1.8	-0.24	0.15	–
Construction	3.0	11.8	1.3	0.05	0.11	–	-1.1	7.7	2.9	0.17	-0.07	–
Infrastructure services	-0.4	3.5	-3.6	-0.57	-0.06	–	1.5	5.7	-1.7	-0.30	0.26	–
Education & health	-0.1	6.3	-1.1	-0.04	0.00	–	1.0	3.8	-3.1	-0.13	0.04	–
Energy & natural resources	-1.2	3.1	-3.1	-0.03	-0.01	–	-1.7	-0.1	-4.3	-0.04	-0.02	–
Other services	-0.8	15.2	8.6	0.45	-0.04	–	0.1	2.5	-3.5	-0.27	0.01	–
GDP (overall economy)	0.0	**2.63**	-4.8	**−3.01**	**−1.78**	**7.5**	0.0	**2.7**	-3.2	**−2.24**	**−0.96**	**5.90**

Source: Authors' calculations from data in GSO (2014)

investment bias. Further analysis is required to disentangle the contributions to employment growth from each potential explanation.

From 2000 to 2007 employment growth is shown to be 4.8 per cent per year less than GDP growth, with reductions in labour use (gaLi) contributing 3.01 per cent and sectoral share changes 1.78 per cent. Hence, biased technological change and/or institutional wage bias are responsible for 63 per cent of the difference between GDP and employment growth, and structural transformation and/or state investment bias are responsible for the remaining 37 per cent. From 2007 to 2011, the difference between GDP and employment growth was 3.2 per cent, with labour productivity improvement contributing 2.24 per cent (70 per cent of this difference) and structural shifts adding 0.96 per cent (30 per cent of the difference). These results suggest that structural transformation explains at most about one-third of the growth rate differential.

The key sectors determining slower employment growth are agriculture, manufacturing, infrastructure services and construction. Because of its size and low labour productivity, changes in the traditional agricultural sector are the greatest factor in these calculations. Agriculture's sectoral output share fell at 3.7 per cent per year from 2000 to 2007, and labour productivity improved at 4.1 per cent per year, resulting in employment growth rate reductions of 2.2 and 2.4 per cent, respectively. From 2007 to 2011 these contributions were 1.34 and 1.42 per cent, once again with labour productivity improvements being a slightly greater factor than the declining share of agricultural output. For manufacturing, in each period increases in employment due to expanding sector shares are more than offset by improved labour productivity. That improvement was much faster, at 4.6 per cent per year in the earlier period than at 1.8 per cent in the latter period. For infrastructure services there is a mixed sectoral share story but productivity improvements during each period, with output share increases compensating for productivity improvement only in 2007–2011. Construction is the one sector for which labour productivity declined, at 1.3 per cent per year from 2000 to 2007 and 2.9 per cent per year from 2007 to 2011. Its output share increase in 2000–2007 (3.0 per cent) increased employment, but its output share decline from 2007 to 2011 (1.1 per cent) partially offsets the increasing demand for labour in this sector. Energy and natural resources take a large share of investment and output, mostly by state-owned enterprises, and that share is also shrinking. At the same time, the very small employment share for this sector makes its contribution to employment growth minimal. Other

services, including professional and various public services, show declining sector shares from 2000 to 2007, and improving labour productivity after 2007.

4 Technological Change and Employment

One issue in interpreting the above employment and growth decomposition is to identify whether increasing labour productivity is because of capital-labour substitution or due to technical change. Moreover, technical change can be labour augmenting if productivity of labour grows faster than productivity of capital. While previous studies generally assumed Hicks-neutral technological change, critics of these studies point out that ignoring biased technological change may lead to underestimation of TFP growth (for example, Rodrik, 1997).

Estimation of substitution elasticities and of changing factor productivity (biased technical progress) is challenging because of the nature of the data used to estimate production functions. Figure 2 shows labour and capital productivity from 2001 to 2011 for the overall Vietnamese economy. It also shows real private wages, real interest rates and thus the wage-rental ratio over that same period. While rising wages and steadily rising labour productivity are evident, fluctuating real interest rates led to an erratic wage-rental ratio that rose significantly at the beginning of this period and is confused by negative real interest rates at the time of the global recession. While capital productivity was stagnant until 2008, it fell dramatically during and after the great recession, even as the wage-rental ratio moved back to its pre-recession levels. If these volatile relative factor prices are driving factor substitution, it is not evident in this data. More likely, long-run expectations of rising real wages may have helped the Vietnam transition to more efficient, albeit capital-intensive technologies. Assuming constant technical progress better fits this medium-term data, not only for the overall economy, but for most disaggregated sectors as well.

Econometric estimation of production functions using this type of data often leads to estimated elasticities of substitution not significantly different from zero and significantly different from one. Disentangling the elasticity of substitution estimate from biased technical progress requires an identifying assumption, typically a choice of functional form. We examine three alternatives in the set of CES models—Cobb-Douglas, Leontief and a general CES specification. In addition, it has been recognized that

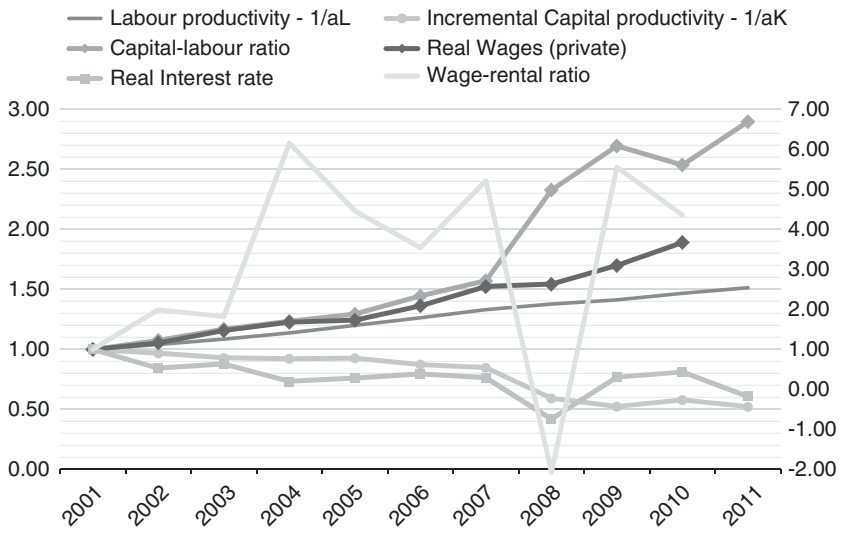

Fig. 2 Factor prices, substitution and technical change in Vietnam. (Notes: All variables are expressed as indices normalized to equal 1 in 2001. Sources: GSO (2011, 2014), IMF (2013))

investment and growth may be determined simultaneously, introducing an identification problem (Levinsohn & Petrin, 2003). Estimation of factor demand equations, one of several approaches used to address these problems, may address that concern at least for labour demand. Estimation of disaggregated sectoral factor demands also reduces the likelihood of simultaneity in capital demand equations.

Previous studies often used an accounting approach to calculate TFP growth under a Cobb–Douglas production function to avoid these same problems, and to identify TFP growth. In that approach, actual input cost shares are employed to measure the parameters of the production function. The broad results based on an accounting approach often seem reasonable, but they may hide potential weaknesses associated with the assumptions of the Cobb-Douglas production function. That production function is believed to be robust if similar results can be obtained from both accounting and econometric approaches. Results from econometric estimation can also test whether elasticities of substitution are indeed one,

but must separately identify technical progress and factor substitutions because of changes in the wage-rental ratio.

To circumvent potential problems associated with biased technological change and the validity of the production functional forms assumed, the analyst can estimate productivity growth rates in factor demand equations, allowing both Hicks-neutral and factor-augmenting technological change, derived from Cobb-Douglas, CES and Leontief production functions, and using both accounting and econometric approaches. In our case, we use sectoral data from GSO (2009, 2010) to estimate productivity growth rates for 18 aggregated sectors and for the overall economy from 2000 to 2008. Detailed derivations and results are available in Abbott et al. (2011). In this chapter, we report key results in Table 4 and discuss the lessons learned.

Accounting approaches based on the Cobb-Douglas production function are clearly the most common in the literature, and their application to the Vietnamese data leads to reasonable though somewhat low estimates of Hicks-neutral technical change, similar to results found for other Asian countries and by ILSSA (2008) for Vietnam. The econometric estimation of a Cobb–Douglas production function gives vastly different results for parameters that should equal costs shares, and in several cases these parameters are negative unless constrained at zero. The estimates of technical progress are only somewhat different from accounting results. Specifications such as these do not predict factor-use well over the sample period.

The assumption of a unitary elasticity of substitution must be relaxed to examine biased technical change. Our attempts to apply a general CES production function provide evidence of quite low elasticities of substitution, but results are not robust and some sectoral parameters are found in implausible ranges—reflecting the data concerns noted above. Moreover, estimated elasticities of substitution are much smaller when biased technical progress is taken into account, as significant labour-augmenting technical progress is then found and better explains output and employment. Estimation using a Leontief production function, hence constraining the elasticity of substitution to equal zero, provides the most robust results. Estimates of TFP growth are not unreasonably large, but larger than found using the more standard approaches. For the overall economy

Table 4 Estimated sectoral elasticities of substitution and technical progress

Sector	*Elasticities of substitution*		*Technical progress (% per year for labour efficiency improvement)*				
	CES estimates		*Cobb-Douglas*		*CES estimates*		
	Neutral	*Biased*	*Accounting*	*Estimated*	*Neutral*	*Biased*	*Leontief*
Agriculture	-0.38	0.13*	5.4	7.6	82.0	7.0	5.7*
	(-0.86)	(3.34)	–	(1.2)	–	–	(16.5)
Manufacturing	1.52*	0.87*	3.4	4.6	2.0	5.2	6.0*
	(24.9)	(2.35)	–	(1.5)	–	–	(22.8)
Construction	0.28	0.69*	4.3	4.0	5.3	10.1	5.8*
	(1.41)	(10.70)	–	(1.9)	–	–	(3.6)
Infrastructure services	0.4*	0.24	9.1	18.0	1.4	15.7	12.9*
Education & health	1.04*	0.71*	7.5	7.6	0.7	28.0	9.0*
Energy & natural resources	0.77*	0.86*	-0.7	-0.4	-0.2	-2.7	0.6
GDP (overall economy)	1.54*	0.36*	4.1	2.8	4.5	10.4	5.8*
	(2.31)	(5.04)	–	(1.7)	–	–	(27.6)

Notes: Estimates assuming Cobb-Douglas and CES production functions yield Hicks-neutral technical change. Biased technical change estimates using CES and Leontief production functions yield separate labour and capital efficiency improvements. Only labour results are presented here. For infrastructure services, education and health, and energy and natural resources, estimates reported are averages over several more disaggregated subsectors. *T*-statistics are reported in parentheses below estimates where applicable. Where averages are reported, for the accounting estimates, and for technical progress in the CES cases where measures are derived, *t*-statistics cannot be easily estimated and thus are not reported. Statistical significance (at 1 per cent) is indicted by * where *t*-statistics can be computed and for the averages when there are consistent results across all subsectors

Source: Abbott et al. (2011)

labour efficiency is found to grow at 5 per cent per year, while capital efficiency grows at 1 per cent per year.[3]

In sum, comparisons of the predictive ability of alternative production functions and thus derived conditional employment demand specifications

[3] Detailed sector data provided by GSO (2009) are not entirely consistent with the aggregate data on its website, with the largest differences in deflated capital allocations. Hence, the aggregate data show slower labour productivity growth and declining capital productivity. Gross output rather than value added was also used in these estimations, and thus productivity differences may also reflect differences in intermediate input efficiency.

confirm the robustness of the Leontief production function results, given data constraints. These results are consistent with significant labour-augmenting technical progress and better explain recent economic performance in Vietnam. Hence, the most convincing interpretation of the labour demand growth decomposition is based on the Leontief results that suggest that increasing labour productivity is because of biased technical progress, not capital-labour substitution.

5 Institutional Biases and Employment

In Vietnam there are two policies that may have biased development towards a capital-intensive path: minimum wage policy and state investment policy. In this section, we investigate to what extent these two policies induce capital-intensive production, and how they in turn affect employment in various key economic sectors.

Minimum Wage Policy

If minimum wage policy is not an important determinant of market wages, which rather follow inflation and productivity trends, then the role of the minimum wage in employment over the medium run is minimal. The literature on the magnitude and direction of these effects in Vietnam and elsewhere is conflicting and unresolved (Cuong, 2010). There is also controversy over what data is appropriate, with some emphasizing the use of only micro data. But both Cuong (2010) and Sakellariou and Fang (2014), who look mostly at distributional implications over the longer term, lack sufficient variation in cross-sectional or panel household data (the VHLSS survey) to identify significant effects directly attributable to minimum wages. These studies also note that the minimum wage is not always respected, not only by informal firms but also by some formal firms. A long-run effect noted by McCaig and Pavcnik (2013) is that firms may be reluctant to register as formal enterprises because of minimum wage laws. Yet this effect is muted if formal firms are paying only a small share of workers at or below the minimum wage.

Here we focus on whether minimum wages distort capital-labour ratios. We first explore whether the minimum wage is seen as an important factor in setting overall market wages. Then we consider to what extent the minimum wage would affect labour-output coefficients and employment, using the conditional employment demand functions discussed in the

section 'Technological change and employment'. If elasticities of substitution are near zero (our earlier result), even if minimum wages do drive overall wages, the effect on labour intensity and thus employment is small.

Ideally, we would also compare minimum wages to wages of unskilled workers. Unfortunately, available data do not allow such distinctions. If the minimum wage is binding, it is most likely for only a small number of low-skilled workers in a firm who represent only a fraction of the workforce. The question we can address is whether available evidence supports the claim that minimum wages influence the overall wage structure, and hence average wages.

Figure 3 compares wages for state-owned enterprises, foreign-invested enterprises and the private (formal) sector to the minimum wage from

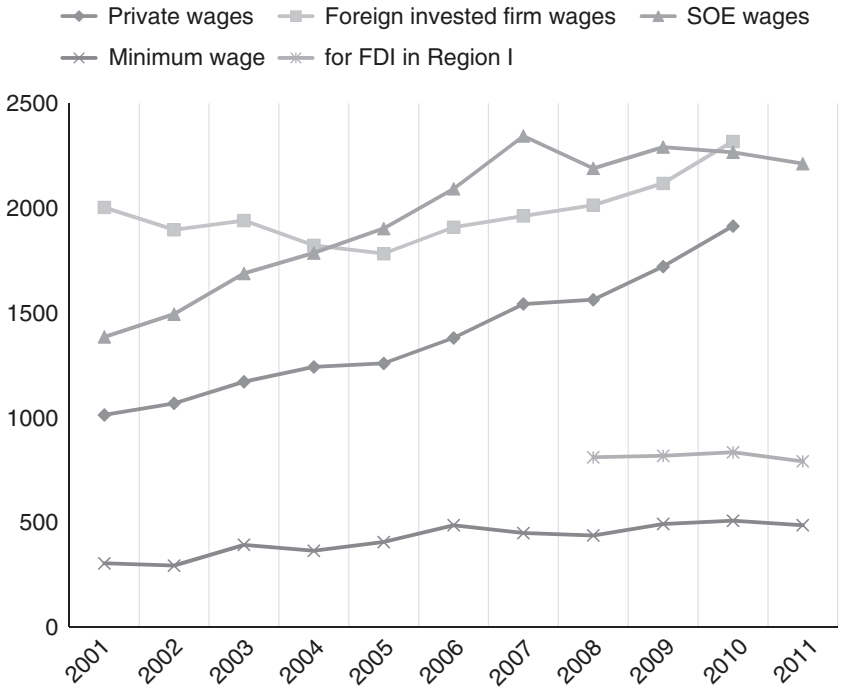

Fig. 3 Real wages and the minimum wage in Vietnam. (Notes: Real wages are monthly average income per employee per month in thousand dong (at 2005 prices). Sources: GSO (2011, 2014), Wikipedia (2013), IMF (2013))

2001 to 2011. Two observations are apparent. First, the minimum wage is much lower than actual average wages for any of these firm types. After 2007 there were separate minimum wages for private firms and for foreign-invested firms. Average wages for private formal firms were subsequently nearly four times the minimum wage applicable to those firms, and wages in foreign-invested firms were nearly three times higher than the much higher minimum wage required of those firms. While wages for informal firms and traditional labour markets are lower, the minimum wage is seldom enforced in those cases.

The second observation is that changes in minimum wages often lag rather than lead wages paid by the various firm types. Once set, minimum wages have often not been changed for more than a year. From 2000 to 2011 the minimum wage was changed eight times, so that sometimes the lag has been as long as two years. As seen in Fig. 3, private wages have been steadily increasing in both nominal and real terms, as both inflation and economic growth have fuelled pressure for higher wages. For state-owned enterprises minimum wages have changed more frequently since 2007, but have remained flat in real terms, as have the levels of real state wages. In private (formal and informal) and foreign-invested firms, however, real wages have increased rapidly in periods when real minimum wages remained constant or fell.

To measure the link between minimum wages and actual nominal wages, we estimated wage determination models. They included the real minimum wage and the CPI as independent variables to explain variation in the nominal market wage. The goal of this regression is to determine the importance of the real minimum wage and/or inflation in determining the overall pattern and level of nominal market wages. We estimated this relationship separately for each firm type and for each aggregated sector. Private (formal sector) wage data as well as wages for foreign-invested firms came from the Enterprise survey (GSO, 2011). The GSO website (GSO, 2014) reports wage data for state-owned enterprises, from which more current data are available. CPI values are from IMF (2013) and minimum wage data based on the Labour Law by the National Assembly of Vietnam that are accessible online (Wikipedia, 2013).

Estimation results summarized in Table 5 indicate that for most sectors inflation is a significantly more important factor than the minimum wage in driving market wages. All of the elasticities with respect to the real minimum wage are small and are statistically significant in only 3 of 12 cases. We also tried constraining the elasticity on inflation to one, that is, neutral

Table 5 Estimated coefficients on the real minimum wage and CPI in sectoral nominal wage determination models

Sector	*Firm ownership*					
	State		*Private*		*Foreign*	
	real min wage	*CPI*	*real min wage*	*CPI*	*real min wage*	*CPI*
Agriculture	1.02*	1.99***	2.01*	-0.35	0.41	1.10***
	-3.17	-8.91	-2.89	(-0.72)	-1.53	-7.84
Manufacturing	0.34	1.43***	0.16	1.46***	0.3	0.89***
	-2.25	-8.14	-1.02	-8.26	-2.25	-11.05
Construction	0.46*	1.27***	0.43	1.50***	-0.064	1.54***
	-2.61	-9.11	-1.42	-6.3	(-0.14)	-7.56
Infrastructure	0.21	2.13***	0.037	1.72***	0.03	1.05***
services	-1.36	-16.94	-0.36	-20.6	-0.25	-18.88
Overall—All sectors	0.59*	1.61***	0.28	1.60***	0.3	0.95***
	-2.89	-9.39	-1.59	-10.8	-1.73	-11.71

Notes: Numbers in parentheses are *t*-statistics. * represents $p < 0.05$, ** represents $p < 0.01$ and *** represents $p < 0.001$

Dependent variable is the nominal sectoral wage

Source: Wu (2012)

inflation pass-through for all the sectors. Then only five cases have a significant (but low) correlation between the real market wage and the real minimum wage. For the original model (Table 5), two of the three sectoral cases where a significant relationship was found were for state-owned enterprises. The other case was private invested agricultural firms. For the economy overall, coefficients are insignificant for private and foreign-invested firms, while that for state-owned firms is only significant at the 5 per cent level. On the other hand, inflation is a highly significant factor determining nominal wages, at the 0.1 per cent level in 11 of 12 sectoral cases and also overall (independent of ownership). Hence, this evidence suggests minimum wages have had little influence on the evolution of market wages in Vietnam. If it had a small effect, it was on state-owned firms and not private or foreign-invested firms. Moreover, since some sectoral coefficients on inflation differ from one, there appears to be some degree of non-neutrality in inflation pass-through.

The second factor in the link between minimum wages and employment is the extent to which higher wages would induce factor substitution

towards more capital-intensive techniques. Our estimates of elasticities of substitution yielded quite low, statistically insignificant values. Had minimum wages significantly increased market wages, then their effect on employment would have been muted by these small elasticities. The combination of minimum wages exerting little influence on overall market wages plus these very limited effects of wages on factor demand patterns suggests that the recent evolution of minimum wages did not contribute in any significant way to more capital-intensive development. Thus, the improved labour productivity term in equation (3) should be attributed to biased technical progress, not to this institutional bias.

State Investment

According to MOLISA and ILO (2010), investment in state-owned firms tends to be overly capital intensive, and they could absorb more labour if that investment were less so, or if more investment went to the private sector. In order to evaluate the role of state investment in recent employment trends, two issues must be addressed. One is the extent to which state-owned enterprises are more capital intensive than other firms. The second is the extent to which there has been restructuring, resulting in either a smaller state sector or state enterprises concentrated in capital-intensive sectors. This analysis is limited by the data available. The GSO (2014) website reports output, investment and employment by either firm ownership type or sector. It does not provide tables that address both these dimensions at once. This limits both capital intensity and restructuring information.

Figure 4 presents both labour and capital productivity by firm ownership. It is apparent that private domestic firms use much less capital and much more labour per unit output (value added) than either foreign-invested or state-owned firms. As of 2011, capital and labour intensities of state-owned and foreign-invested firms were similar. Labour productivity in state-owned firms at that time was six times higher than in domestic private firms, while capital productivity was less than one half that for private firms. Both state-owned and foreign-invested enterprises clearly chose more capital-intensive techniques than did domestic private firms.

These factor intensities suggest that investments by state-owned enterprises go counter to Vietnam's comparative advantage as a labour-abundant, low-wage country. This reflects that these firms have had access to better, more capital-intensive international technologies than do strictly

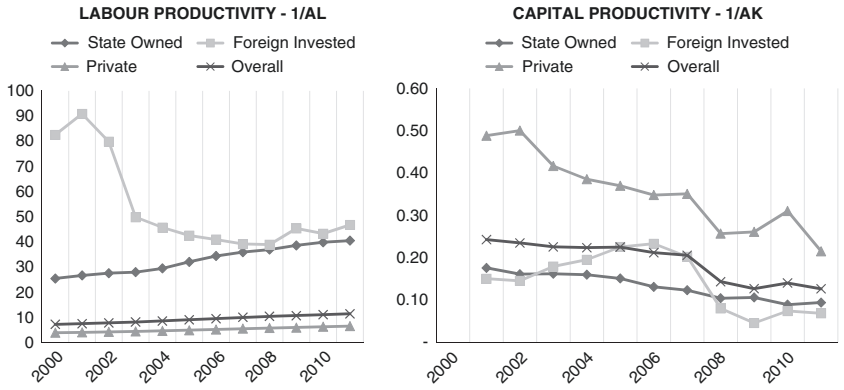

Fig. 4 Labour and capital productivity by firm ownership. (Notes: Real value added per worker is measured in million dong (at 1994 prices) per year. Capital productivity is value added per year in million dong per million dong invested. Source: GSO (2014))

private firms, especially as state-owned enterprises have in the past partnered with foreign firms. Allocation of savings by a socialist government that favours state-owned enterprises has therefore acted as an implicit capital subsidy.

Factor intensities have changed over time because of both shifts in sectoral emphasis and technical change. Foreign-invested firms were in 2001 much less labour intensive than even state-owned firms. Since then they have invested in more labour-intensive sectors, and thus employment intensity has increased. Capital intensity for foreign-invested firms initially decreased as employment intensity increased, but there was a significant decline in capital productivity for all ownership types after 2007. While labour productivity has been steadily increasing for all firm types over the entire period for state-owned and private firms, capital productivity was declining, corresponding with labour-capital substitution or biased technical change. The fall in capital productivity after 2007 was not matched by any change in the trend for labour productivity. If there is an effect on employment trends due to capital-intensive development, it appears to be found for all firm types, and the changing role of foreign-invested firms is the most dramatic change.

Table 6 presents GDP, investment and employment by firm ownership in 2000, 2007 and 2011 to show the extent of restructuring by

Table 6 Investment, output and employment by firm ownership

	Overall	*Enterprises by firm ownership*		
		State owned (SOE)	*Foreign invested (FDI)*	*Private*
2000				
GDP	274	112	30	133
		41%	11%	48%
Investment	115	68	21	26
		59%	18%	23%
Employment	37.1	4.4	0.4	32.4
		12%	1%	87%
2007				
GDP	461	180	61	220
		39%	13%	48%
Investment	309	132	85	93
		43%	27%	30%
Employment	45.2	5.0	1.6	38.7
		11%	3%	86%
2011				
GDP	584	213	80	291
		37%	14%	50%
Investment	363	145	95	122
		40%	26%	34%
Employment	50.4	5.3	1.7	43.4
		10%	3%	86%

Notes: GDP and investment are in trillion dong. Employment is in million persons. Shares are reported in per cent below each variable

Source: GSO (2014)

state-owned enterprises as well as the emergence of foreign-invested firms. These data suggest very little restructuring, but reflect a state sector being slowly overtaken by more dynamic private and foreign firms. The state share of GDP fell from 41 per cent in 2000 to 39 per cent in 2007 and 37 per cent in 2011. The decline in its share of investment was much larger, from 59 per cent in 2000 to 40 per cent in 2011. Investment declines greater than output declines reflect increasingly capital-intensive production and sectoral shifts, as well as declining capital productivity, for all types of firms. The employment share of state-owned enterprises changed very little over this period, at 3 per cent in both 2000 and 2011.

With this small change in employment shares, state enterprises play a very limited role in our employment growth decomposition. Had there

been much more restructuring, the differences in capital and labour intensity would have mattered more. Possibly more notable are the changes for foreign-invested firms. Their investment share increased from 18 per cent in 2000 to 26 per cent in 2011 following the surge in foreign investment around WTO accession in 2008. The investment share of private domestic firms increased faster than for either state or foreign firms, from 23 per cent in 2000 to 34 per cent in 2011. The employment share of foreign firms jumped from 1 per cent in 2000 to 3 per cent in 2007 and 2011. Their output share increased somewhat from 11 per cent in 2000 to 14 per cent in 2011. The small share changes are all because of the dynamic growth of private firms that makes changes in state and foreign firms relatively smaller.

Sector composition matters to capital intensity and to the scope for restructuring. State-owned enterprises have been heavily involved in capital-intensive mining, natural resources and energy sectors, although they operate in many other, potentially more labour-intensive sectors as well. Foreign firms have also exhibited some significant sectoral shifts since 2000. Revised investment and enterprise laws in the mid-2000s, as part of pre-WTO accession changes, allowed wholly owned foreign-invested firms that are no longer required to partner with SOEs to alter their sectoral composition. This led to an increase in the share of foreign firms in manufacturing. Consequently, the state's share in manufacturing declined.

Older data are available by sector and firm ownership for investment (but not output) (GSO, 2009), allowing us some idea of what sectoral shifts have been occurring as shown in Table 7. This table shows the large decline in the overall investment share of state-owned enterprises, but not dramatic changes in the sectoral composition of that investment allocation. The agriculture share dropped the most while construction increased, and there is a decline in the share of the energy and natural resource sectors matched by an increase for foreign firms, similar to that for manufacturing. This does not provide a consistent story about SOE movements into or out of capital-intensive sectors, as some expanding sectors were capital intensive while others are labour intensive. Moreover, movement of investment out of agriculture has only limited structural transformation effects, as it was for processing and distribution firms, not primary production where most low-wage employment is found.

These results suggest that restructuring and investment allocation into the private firms, which are relatively less capital intensive, could be an important factor driving future employment growth. If state firms had

Table 7 Sectoral investment shares by firm ownership

Sector	*Firm ownership (%)*					
	State owned (SOE)		*Foreign invested (FDI)*		*Private*	
	2000	*2007*	*2000*	*2007*	*2000*	*2007*
Agriculture	4.6	2	3.4	2.4	4.8	4
Manufacturing	4.8	3.2	6.8	9.3	3.1	3.4
Construction	1.4	1.8	0.5	0.3	0.5	1.3
Infrastructure services	12.7	8.7	0.9	2	1	5.4
Education & health	9.3	6.9	0.6	0.5	0.6	0.9
Energy & natural resources	16.2	9.3	1	8.2	0.5	1.8
Other services	10.1	8	4.8	2.1	12.3	12.6
GDP (overall economy)	59.1	39.9	18	24.8	22.9	29.5

Source: GSO (2009)

been expanding in output shares, then the employment growth rates in those modern sectors would be lower because of relatively low labour requirements in state enterprises. This story is probably more relevant to the 1990s than to the 2000s, however. The slowly declining importance of state production in key modern sectors weakens the impact of an overly capital-intensive production strategy pursued by state firms on slowing down employment growth for the entire economy. The low state employment share, stagnant output share and limited restructuring mean these sectoral ownership effects are a small part of our employment growth decomposition. At the same time, it may be the case that restructuring could occur more rapidly in the future, making the gap between GDP and employment growth smaller as private firms expand more rapidly.

6 Conclusions

Employment in Vietnam has grown more slowly than GDP over the last two decades. While this difference reflects improvements in labour productivity and rising GDP per capita—signs of successful development—it has also raised concerns among Vietnamese policymakers that economic development is not creating enough new jobs. Structural transformation has moved labour from lower-productivity traditional sectors, especially agriculture, to higher-productivity modern sectors including manufacturing and services. The Vietnamese labour ministry in its assessment of the labour

situation in Vietnam (MOLISA and ILO, 2010) believes this restructuring is not moving sufficiently rapidly, has involved too little innovation, and exhibits an overly capital-intensive development strategy.

We used data for seven aggregate sectors and the overall Vietnamese economy from 2000 to 2007, and then to 2011, provided by GSO (2014), to examine the roles played by structural transformation, technical progress and institutional bias towards capital-intensive development to explain the differences observed between employment and GDP growth. This decomposition attributes only about one-third of the employment-growth difference to shifts from low-productivity to higher-productivity sectors. The remaining two-thirds came from declining labour use per unit output that is also found in agriculture. Further analysis attempted to separate technical progress and structural transformation from institutional biases towards capital-intensive development.

Data on capital- and labour-output ratios reveal steadily improving labour productivity, but erratic and inconsistent factor price trends. Estimated elasticities of substitution are generally significantly less than one, and not significantly greater than zero, regardless of specification. Estimation using a Leontief production function shows that TFP growth is not unreasonably large, and larger than found using the more standard approaches relied on in MOLISA and ILO (2010). For the overall economy, our results suggest that labour efficiency grew at 4.2 per cent per year, while capital efficiency grew at 1 per cent per year. These results are consistent with significant labour-augmenting technical progress and better explain historical economic performance in Vietnam. This is highly consistent with the Rodrik (1997) conjecture that there may have been significant technical progress behind the Asian miracle, and that the analyst must address a labour-augmenting bias in technical change to find it.

Our results also suggested that rising minimum wages only contribute to a very limited extent to more capital-intensive development. We looked at both of the two key links between the minimum wage and labour intensity—the effect of minimum wages on overall wages and the effect of wages on technical choices. Real market wages have been rising for private formal and foreign firms as real minimum wages held constant since 2007. We found that non-neutral inflation pass-through better explains nominal wage evolution than do minimum wages. Therefore, the differential between GDP and employment growth because of falling labour output ratios is mostly because of technical progress. Undoubtedly, long-run expectations on relative wages have played a role in sectoral and technical

choices made in Vietnam. Yet, over the short run institutional biases in wage setting are only a limited factor.

Investment by state-owned enterprises appears to be much more capital intensive than for private firms. While substantial investment in SOEs has persisted, low capital productivity since 2007, low labour output ratios and much more rapid growth of private firms mean that this factor has only played a limited role in recent Vietnamese employment trends. This is best viewed as a missed opportunity. Since there was little restructuring, there is little evident impact on employment. More rapid restructuring and greater investment allocations to private firms and less to SOEs would likely lead to greater demand for labour.

Structural transformation only partially explains the employment growth seen in Vietnam. While some of the difference from GDP growth may, as just noted, be attributed to capital-intensive investment by the state, a significant share of the employment-growth difference is because of technical change when low elasticities of substitution and labour-augmenting bias in technical change are taken into account. In sum, there is little reason to be pessimistic about the performance of the Vietnamese economy. Focus in the coming years should be on consolidating the distinct progress made rather than on promoting policy shifts based on inadequate interpretation of the available evidence.

Acknowledgements The authors appreciate helpful comments from conference participants at the Central Institute for Economic Management (CIEM), Hanoi, Vietnam in June 2010 and March 2011, and the Asian Development Review Conference (ADR) held at the Asian Development Bank (ADB) in Manila on 1–2 August 2013. Thanks are also due to Simon McCoy for advice and collaboration, and to Danida for financial support. Furthermore, we are grateful to two anonymous referees for constructive critique and to the EJDR editor for guidance. The usual caveats apply.

References

Abbott, P. C., & Tarp, F. (2012). Globalization Crises, Trade and Development in Vietnam. *Journal of International Commerce and Economic Policy, 3*(1), 1–23.

Abbott, P. C., Tarp, F., & Wu, C. (2011). *Structural Transformation, Biased Technological Change, and Labour Demand in Vietnam*. Prepared for Central Institute of Economic Management (CIEM), Hanoi, Vietnam, and University of Copenhagen, Copenhagen, Denmark.

Collins, S. M., & Bosworth, B. P. (1996). Economic Growth in East Asia: Accumulation Versus Assimilation. *Brookings Papers on Economic Activity, 2*, 135–203.

Cuong, N. V. (2010). *The Impact of a Minimum Wage Increase on Employment, Wages and Expenditures of Low-Wage Workers in Vietnam*. MPRA Paper No. 36751, Munich, Personal RePEc Archive, December.

Freeman, C. (1995). The 'National System of Innovation' in Historical Perspective. *Cambridge Journal of Economics, 19*(1), 5–24.

General Statistical Office of Vietnam (GSO). (2009). *Investment, Trade and Price Data for Vietnam*. General Statistical Office of Vietnam.

General Statistical Office of Vietnam (GSO). (2010). *Statistical Yearbook of Vietnam*. General Statistical Office of Vietnam (and Earlier Volumes).

General Statistical Office of Vietnam (GSO). (2011). *Enterprise Survey*. General Statistical Office of Vietnam (and Earlier Surveys).

General Statistical Office of Vietnam (GSO). (2014). *Statistical Data Website*. General Statistical Office of Vietnam. Retrieved October 14, 2014, from http://www.gso.gov.vn/default_en.aspx?tabid=491

Institute of Labour Science and Social Affairs (ILSSA). (2008). *Forecasting the Relation between Investment, Economic Growth and Employment and Income of Employees*. Institute of Labour Science and Social Affairs (ILSSA).

International Monetary Fund (IMF). (2013). *International Financial Statistics*. IMF.

Kapsos, S. (2005). *The Employment Intensity of Growth: Trends and Macroeconomic Determinants*. Employment Strategy department, ILO.

Kim, J.-I., & Lau, L. J. (1994). The Sources of Economic Growth of the East Asian Newly Industrialized Countries. *Journal of the Japanese and International Economies, 8*(3), 235–271.

Krüger, J., Cantner, U., & Hanusch, H. (2000). Total Factor Productivity, the East Asian Miracle, and the World Production Frontier. *Review of World Economics, 136*(1), 111–136.

Krugman, P. (1994). The Myth of Asia's Miracle. *Foreign Affairs, 73*(6), 62–78.

Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies, 70*(2), 317–342.

McCaig, B., & Pavcnik, N. (2013). *Moving out of Agriculture: Structural Change in Vietnam*. NBER Working Paper 19616. NBER, Cambridge, MA.

Ministry of Labour, Invalids and Social Affairs (MOLISA) and International Labour Organization (ILO). (2010). *Labour and Social Trends in Viet Nam 2009/10*. Ministry of Labour, Invalids and Social Affairs (MOLISA) and International Labour Organization (ILO).

Nelson, R. R., & Pack, H. (1999). The Asian Miracle and Modern Growth Theory. *The Economic Journal, 109*(457), 416–436.

Pack, H., & Page, J. M., Jr. (1994). Accumulation, Exports, and Growth in the High-performing Asian Economies. *Carnegie-Rochester Conference Series on Public Policy, 40*(1), 199–235.

Rodrik, D. (1997). *TFPG Controversies, Institutions, and Economic Performance in East Asia*. NBER Working Paper No. 5914. NBER, Cambridge, MA.

Rodrik, D. (2015). *Premature Deindustrialization*. NBER Working Paper No. 20935. NBER, Cambridge, MA.

Sakellariou, C., & Fang, Z. (2014). The Vietnam Reforms, Change in Wage Inequality and the Role of the Minimum Wage. *Economics of Transition, 22*(2), 313–340.

Wikipedia. (2013). The Minimum Wage in Vietnam. Retrieved October 14, 2014, from http://translate.google.com/translate?js=y&prev=_t&hl=zh-CN&ie=UTF-8&layout=1&eotf=1&u=http://vi.wikipedia.org/wiki/L25C625B025C625A1ng_t25E125BB2591i_thi25E125BB2583u_t25E125BA25A1i_Vi25E125BB2587t_Nam&sl=vi&tl=en

World Bank. (1993). *The East Asian Miracle: Economic Growth and Public Policy: Summary*. Oxford University Press.

World Bank. (2013). *World Development Indicators (WDI)*. World Bank. http://data.worldbank.org/indicator

Wu, C. (2012). *The Myth and Mystery of Vietnam's Economic and Labour Demand Growth* (Ph.D. Dissertation). Purdue University, West Lafayette, IN.

Young, A. (1995). The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience. *The Quarterly Journal of Economics, 110*(3), 641–680.

‘Helping a Large Number of People Become a Little Less Poor’: The Logic of Survival Entrepreneurs

Erhard Berner, Georgina M. Gomez, and Peter Knorringa

1 Introduction

Almost since the discovery of the ‘informal sector’ of urban economies in the 1970s, researchers have noticed the existence of a subcategory that faces particular barriers to growth, and has been termed ‘street economy’, ‘(sub-)subsistence production’, ‘necessity-driven entrepreneurs’, ‘informal proletariat’, or ‘survival(ist) enterprises’. However, a deeper economic and sociological analysis of this group’s specific characteristics is largely missing. This chapter seeks to characterize the different rationalities of the various categories of entrepreneurs in order to better understand their behaviour and distinct needs. We adopt the concepts of ‘survival’ and ‘growth-oriented’ entrepreneurs to reconstruct the respective logics of

E. Berner (✉) • G. M. Gomez • P. Knorringa
Institute of Social Studies, Erasmus University Rotterdam, The Hague, The Netherlands
e-mail: berner@iss.nl; gomez@iss.nl; knorringa@iss.nl

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_17

their strategies. We will stress the importance of social relationships in determining their objectives, motivations, and preferences. Our focus is on the different logics of entrepreneurship in the informal economy, and we aim at reflecting on these differences in order to link groups of entrepreneurs and policy interventions in a more consistent manner. Other researchers like Mead and Liedholm (1998) and Rogerson (1997) have also identified these two distinct groups and chosen to focus on growth-oriented entrepreneurs. Our complementary focus in this chapter is on survival entrepreneurs.

The next section sets the scene by means of an 'anthropological slum walk' to emphasize the vibrancy of economic activity among the urban poor and the moral economy in which these entrepreneurs are embedded. In Sect. "Differences Among Entrepreneurs: A Revisit of the Literature" we follow the trail of the various academic contributions discovering a duality among poorer entrepreneurs, usually without much reference to earlier accounts. This leads us to propose a consolidated typology in Sect. "A Consolidated Typology of Entrepreneurs", based on a re-interpretation of the literature in combination with our own observations in the field. Our main argument is that there is a fundamental and qualitative difference between survival and growth-oriented enterprises with regard to a number of key variables. Section "The Elusive Mirage of Graduation" reviews the empirical evidence of graduation, the process through which survival enterprises are supposed to expand and transform into growth-oriented enterprises. The concluding section demonstrates that much of the intervention practice still attempts to push survival entrepreneurs towards expansion, resulting in inefficiency and unsustainability. We argue for a more realistic and consistent link between business development policies and the rationales of the two groups of entrepreneurs, and we outline a policy approach that fits the logic of survival entrepreneurs.

2 The Moral Economy of the Slum

Few visitors to poor urban neighbourhoods fail to notice the tremendous intensity and vibrancy of economic activity. Houses, yards, alleys, and streets are sites of construction, production, service provision and trade, a seedbed of entrepreneurship. Yet on closer inspection a lot of these businesses seem to defy the basic logic of entrepreneurship: to invest available capital, apply specialized skills, and make a profit while accepting a risk. Demand for services is limited because most households have more underutilized labour at their disposal than they have money to pay for someone

to cut their hair or wash their clothes. The range of commodities on offer in the numerous shops is rather narrow, such as snacks, soft drinks, alcohol, cigarettes, toothpaste, detergents, and other goods of daily need. Typically, every seventh building along a street houses a store like this, limiting the prospective customer base to some ten families, who in addition patronize open markets and discounters for bulk purchases. What looks like a desperate situation in business terms is clearly quite attractive to slum dwellers. In Manila we asked how they would use a $1000 lottery win, and the majority responded that they would open up yet another neighbourhood store (Berner, 1997, p. 83). Are the poor wasting their precious time and money?

The explanation for this apparently 'irrational' behaviour has several dimensions. First, a truly international proverb provides a rationale for the proliferation of these shops; never put all your eggs in one basket. The store is in most cases part of a strategy of increasing security through diversification, and profits are only one part of the household income, while other sources such as formal or informal wage labour are more often irregular and insecure. The shop brings in inadequate but steady earnings that help to bridge more difficult periods. Moreover, the stock is a form of saving that can be consumed in case of emergency, or sold at buying (i.e. supermarket) price to a 'competitor'. In most cases the shops are run by mothers of young children, elderly persons, or anyone else who has no chance to earn money in another way—people with so-called zero opportunity costs. Finally, many of the sales are on credit. Store owners allow customers to postpone payment to the next payday, and know them well enough to calculate the risk of default.

Among the most baffling observations in slum shops are the ubiquitous strings of portion-packed items such as shampoo or detergent, and the habit of buying cigarettes and chewing gum in singles or pairs while it would be ultimately cheaper to buy a pack. The obvious explanation—poor people just do not have enough money at a given point in time to buy a standard pack or bottle—does not hold water for several reasons. First, a poor person would hardly pay a unit price that is two or three times higher than necessary. In addition, the total sum could be paid later if, as in most cases, the purchase is done on credit anyway. However, the purchase of larger quantities exposes the buyer to the desires of others; a woman who takes a bottle of shampoo or a pack of detergent to a public faucet will have to share with relatives, friends, and neighbours. Buying single-portion packs makes economic sense as users can reserve at least small luxuries—but not for instance staple food—for themselves. It is a

culturally acceptable way of circumventing the powerful obligation to share, exposed in James Scott's seminal *The moral economy of the peasant* (1976), that is prevalent among the poor.

The implications of the moral economy of the slum for entrepreneurs go far beyond petty saving strategies. On the one hand, reciprocity is an essential mechanism to cope with volatility and vulnerability. The life of the poor can be best described as a series of crises, and so being able to turn to somebody in an emergency can literally make the difference between life and death. On the other hand, it also creates a situation that prevents small businesses to accumulate. Once an entrepreneur is more successful than others, that person is expected to pay for the medicine for an aunt's sick child, give loans to customers with questionable creditworthiness, or employ a nephew who is neither capable nor willing to work hard. The aspiring entrepreneur is then likely to slip back into his or her own next crisis, remaining caught in the web of shared poverty.

The permanence in poverty of these risk-avoiding copycats begs the question of whether the concept of entrepreneurship really applies to them. Indeed, they do not fit the model of a Schumpeterian entrepreneur who introduces path-breaking, truly new combinations of factors of production which command a decisive cost or quality advantage. But even in the mainstream economy, very few entrepreneurs engage in such catalytic events which open up a whole new range of economic opportunities (Binks & Vale, 1994). By far, most entrepreneurs in all types of firms, countries, and sectors undertake larger or smaller variations on such initial catalytic ideas. In turn, the entrepreneurs in the slums risk their own assets and independently allocate factors of production, apply their knowledge to their business and make decisions about stocking, changes in their offer, and contracting credit. So, they are entrepreneurs indeed—but, as we will argue in the following, a different type of entrepreneurs with distinct support needs.

3 Differences Among Entrepreneurs: A Revisit of the Literature

A qualitative distinction between various categories of entrepreneurs has been noticed by different scholars, but rarely in relation to the rationale that drives the business. This section recaptures previous theorizations of the two groups of entrepreneurs.

Since the discovery, or invention, of the 'informal sector' by the International Labour Office in 1972 a dualistic view prevailed in analyses of urban economies. The analytical strength of the distinction has been challenged from the start, leading to an endless debate on the definition and operationalization of the terms 'formal' and 'informal' (Turnham et al., 1990 for an early overview). We conclude that these do not denote distinct categories but ideal types or poles of a segmented continuum. Purely formal or informal businesses do probably hardly exist; relations with workers and buyers may be characterized by different degrees of formality; and entrepreneurs react to changes in the regulatory environment by various strategies of formalization and informalization. Moreover, degrees of (in)formality do not provide much analytical information in terms of predicting performance.

Assessment likewise varied sharply from the beginning of the debate. The ILO optimistically stated that 'the bulk of employment in the informal sector far from being only marginally productive is economically efficient and profit making' (1972, p. 5). In contrast, Colin Leys observed that 'the informal sector is in fact a euphemism for cheap labour employment, based on landlessness and unemployment. … it denotes primarily a system of very intense exploitation of labour, with very low wages and often very long hours' (1973, p. 420). Both optimism and pessimism found multiple support in the vast literature of the following decades.

John Friedmann and Flora Sullivan (1974) were the first to propose that both sides could be correct, respectively with regard to one specific segment of a deeply divided informal economy. To them, 'small-scale family enterprises' are able to accumulate some capital 'as a result of competitive advantage and/or superior business acumen' (p. 394), while an irregular or 'street' economy of low-status, low-skilled trading and service delivery offers subsistence-level returns. These findings were echoed by Guy Standing (1977) without reference to Friedmann and Sullivan, the first in an unfortunate tradition of rediscovering and renaming the categories.

Still without dropping informality as distinctive feature, another two-sector model of the informal economy was elaborated by ILO advisor William House in various combinations with other scholars (Rempel & House, 1978; House, 1984; House et al., 1993). In a programmatic definition,

> (t)wo very different groups of people are hypothesized to exist in the informal sector and are distinguished by their activities, attitudes, and motivation. They are labelled 'the community of the poor' and 'the intermediate sector'. Those in the former group are attached to the city in order to gain entrance to employment in the formal sector. (…) They lack the motivation and perhaps the means to seek informal activities with growth potential or to invest in their current activity because they view their situation as temporary. (…) The alternative group in the intermediate sector includes people who have consciously decided on a particular artisan skill or line of business with the intent of making it a means of a more permanent livelihood. Their motivation is to invest and build for the future. (House, 1984, p. 280)

That House's work did not find the deserved recognition may be partly explained by the not very illuminative and appealing terminology. However, his merits lie in the insight that the division is a *qualitative* one, correlated to but not determined by differences in size of investment, number of workers, degree of formality, and so on. Instead, the distinction implies different logics, precisely the point we will reflect on in our further analysis.

Building upon his research in South Africa, Christian Rogerson was the next to observe that 'a useful conceptual distinction can be drawn between two categories of informal enterprise':

> First, are those *survivalist enterprises* which represent a set of activities undertaken by people unable to secure regular wage employment or access to an economic sector of their choice. Generally speaking, the incomes generated from these businesses, which tend to be run by women, usually fall short of even a minimum standard…, with little capital investment, virtually no skills training and only constrained opportunities for expansion into a viable business. Overall, poverty and a desperate attempt to survive are the prime defining features of these enterprises. The second category are *micro-enterprises* or *growth enterprises* which are very small businesses, often involving only the owner, some family members and at most one to four paid employees. These enterprises… have only a limited capital base and their operators possess only rudimentary business skills. Nonetheless, many micro-enterprises have the potential to develop and flourish into larger formal small business enterprises. (1996, p. 171; original emphasis)

Rogerson has meritoriously emphasized the gender dimension of the divide (1996, p. 174, 1997, p. 348f) and raised attention to urban agriculture as an important sector for survival activities. His well-founded

pessimism about the growth potential of survivalist businesses led him to focus on the dynamic growth sector where upgrading and clustering can be pursued (1997, p. 352ff).

In a similar manner, Tellegen (1997) distinguished between 'necessity-driven' and 'opportunity-driven' entrepreneurs, a distinction now used in the Global Entrepreneurship Monitor (Reynolds et al., 2001). This contribution highlights again the differences in motivation to start up a business, apart from the results obtained once in business. While the former are in business to satisfy part of the basic needs of the household, the latter seek to expand the business and improve the living standards of the household at the same time.

Donald Mead and Carl Liedholm, whose unique empirical work we discuss extensively below, also observed the divide: 'Many new and very small MSEs that do not expand in terms of employment are primarily survival-type activities', and thus are particularly appropriate target groups for those concerned with poverty alleviation. These enterprises can be extremely important in 'helping a large number of poor people become a little less poor' (1998, p. 70). However, like Rogerson, Mead and Liedholm chose to focus on supporting more dynamic growth-oriented businesses in their policy recommendations. We take up the challenge to focus on survival entrepreneurs.

A more recent contribution analyses the immediate business system in which enterprises are embedded, apart from the characteristics of the individual entrepreneurs. Survival entrepreneurs operate in an environment characterized by overcrowded market 'niches', negligent or predatory government agents, and multiple but volatile sources of household income—in Geof Wood's (2003) term, by 'destructive uncertainty'. As they have to cope not just with short-term shocks but also predictable hazards, avoiding unnecessary risks is absolutely imperative. At the same time, personalized relations of patronage and reciprocity that offer some security have to be maintained regardless of their long-term costs. To cut off links with exploitative suppliers, buyers, or credit providers, and to break away from the moral economy of sharing, would imply foregoing all claims for emergency assistance. According to Wood, the consequence is a 'Faustian bargain', a discounting of the future in favour of survival in the present, that contributes to chronic poverty: 'The dangers of not being a client, of not being protected, of losing 'membership' of the local commander led community are immense. Better to be with the devil you

know—… *security at the price of graduation*, individual or collective' (2003, p. 468, emphasis added).

From a structural point of view, Alejandro Portes and Kelly Hoffman elaborated a class analysis of the informal economy based on Latin American cases (2003). In line with most previous approaches but again without any reference to them, they propose a bipolar division of the informal sector. The characteristics of the 'petty bourgeoisie' or 'class of microentrepreneurs' (ranked higher than formal workers) are the 'possession of some monetary resources; some professional, technical, or artisanal skills; and the employment of a small number of workers supervised on a direct, face-to-face basis' (Portes & Hoffman, 2003, p. 45). The remainder is defined in a residual way by an absence of these traits, and subsumed to a class of the 'informal proletariat' together with non-contracted wage workers, unpaid family labour, and domestic servants at the bottom of the hierarchy.[1] The number of own-account informal entrepreneurs is in all countries much higher than that of capitalists and petty bourgeoisie combined; earnings are in most cases lower than those of formal workers, and insufficient to sustain a household without additional income sources (Portes & Hoffman, 2003, p. 63).

4 A Consolidated Typology of Entrepreneurs

Distinct categories of entrepreneurs in the informal sector have thus been discovered no less than five times independently. In his *Planet of Slums*, Mike Davis unearthed at least some of this literature and invented 'sub-subsistence' and 'micro-accumulation' as yet another set of labels (2006, p. 180). While all contributions have touched upon elements of the logic of survival business and its difference from 'regular' entrepreneurship, none of them has captured all of them. Most importantly, they have failed to establish a coherent research tradition and consistent, informative terminology—illustrated by the fact that the term 'microenterprise' has been

[1] Portes and Hoffman build on a (mainly Latin American) theoretical tradition that attempted to salvage the 'working class' as emancipatory agent by incorporating 'petty commodity producers'. Indeed work for a fixed wage is only one of various forms of labour relations in the informal sector, including piece-rate homework and 'renting' assets (such as taxis and trucks) from an employer. We see a *qualitative* difference between these ways of organizing dependent labour (mainly found in production and transport) and 'own-account' survival entrepreneurs (typically in trade and petty services).

Table 1 Characteristics of survival and growth-oriented enterprises

Survival(ist)	*Growth(-oriented)*
(Street economy, community of the poor, [microenterprise], necessity-driven, informal own-account proletariat, sub-subsistence)	(Small-scale family enterprise, intermediate sector, [microenterprise], opportunity-driven, petty bourgeoisie, micro-accumulation)
Ease of entry, low capital requirements, skills and technology	Barriers to entry
Female majority	Male majority
Maximizing security, smoothing consumption	Willingness to take risks
Part of diversification strategy, often run by idle labour, with interruptions, and/or part-time	Specialization
Embedded in networks of family and kin	Embedded in business networks
Obligation to share income generated	Ability to accumulate part of the income generated

used for businesses on both sides of the fence, for a third category in between, or as an umbrella term for the whole universe.

The absence of clearly distinct conceptual frameworks as a basis for defined rationales and appropriate performance indicators, leading to a misdirection of scarce resources, has been severely criticized by practitioners (Cotter, 1996; Billing & Downing, 2003). In Table 1 we propose a typology of entrepreneurship[2] as a basis for further research and design of appropriate policy interventions. We have incorporated the contributions from the informal economy literature and enriched it with findings from own research projects across several developing countries.

We do not consider the *scale* of the enterprise and the number of workers it employs as the main distinctive characteristic between survivalist and growth-oriented entrepreneurs. 'Survival enterprise' is not a direct synonym for micro-enterprise. The latter may start with one or two workers, as most survivalists do; some growth-oriented micro-enterprises are able to multiply the number of workers in the business within a few years.

[2] Possible sub-categories within the group of growth-oriented entrepreneurs, for example, distinctions between small, medium, and large businesses, are outside the scope of this chapter. We would assume these to be predominantly quantitative in nature. We also do not go into more detailed sub-categories of survival entrepreneurs, like distinctions between basic survival pre-entrepreneurs and subsistence entrepreneurs, as proposed by Eigen (1992).

These additional employees would be rather specialized or have relevant skills, in contrast to the usually unskilled family members employed out of reciprocal obligation in survival businesses.

In turn, we do not see informality as a necessary common characteristic of both categories. Most survival enterprises would start up and remain within the informal economy, a feature that goes together with the low barriers of entry typical of these enterprises. Growth-oriented micro-enterprises, in contrast, typically start in the informal economy, but often acquire some more formal characteristics when becoming more successful. We would expect almost all of the businesses that significantly formalize at some point in their life to be growth-oriented enterprises. We observed that accessing loans or supplying formal clients often trigger formalization processes, but many growth-oriented enterprises choose not to take this step if such incentives do not exist. In that sense, our analysis of survival versus growth-oriented entrepreneurs crosscuts the formal-informal dichotomy.

It is important to notice that the two columns in the table are not directly juxtaposed, but that regular employment may have an intermediate position in both functional and hierarchical terms. Many survival entrepreneurs would prefer the stability of a regular and stable job, even if at a similar low income as from their enterprise. At the same time, some regular wage workers may accumulate skills and savings that enable them to start a growth-oriented business. Barriers to upward mobility are significant in both instances, so we expect very few people to be able to cross them. A relatively higher likelihood of inter-generational upward mobility, as observed by Barbara Grosh and Gloria Somolekae (1996), supports rather than refutes this argument: Survival businesses cannot lift their owners out of poverty, but may enable some of them to get (continuous) education for their children—a critical precondition for both regular employment and growth-oriented entrepreneurship.

5 The Elusive Mirage of Graduation

Experience shows that the growth potential of survival entrepreneurs is very limited even if they are targeted with well-intended business development programmes. Surprisingly few solid empirical studies on graduation rates exist,[3] in contrast to websites and project documents of development

[3] For a recent literature review on graduation among micro and small businesses, see Gómez (2008).

NGOs and donor agencies that abound with 'success stories' without much information about which type of enterprise did succeed and which did not. The main exception is the Gemini project led by Carl Liedholm and Donald Mead, which systematically collected information on enterprises to empirically test assumptions about survival, death, growth, and graduation. Over a time span of fifteen years starting in 1980, the group gathered data on more than 50,000 enterprises employing up to 50 workers in the Dominican Republic, Botswana, Kenya, Malawi, Swaziland, Zimbabwe, Sierra Leone, Bangladesh, Jamaica, Honduras, Thailand, and Egypt. It was estimated that fewer than 20% of those enterprises with four or less workers grew at all (Mead, 1994, 1999). Even when there was growth, it was only marginal: in the segment of enterprises of four or less workers only 1% graduated to the next size category of more than ten workers (Liedholm & Mead, 1987; Mead & Liedholm, 1998). Other research teams reported similar findings. A study in Mexico found that only 12% of the single-worker firms expanded (Fajnzylber et al., 2006).

Explanations for the absence of growth and graduation can be derived from the characteristics laid out in our typology; survival and growth-oriented entrepreneurs are, in fact, qualitatively different categories. The first one is motivation: Survival entrepreneurs are simply not interested in expanding their business. A survey in eight districts in West Bengal found that the median family had three working members and seven occupations (Banerjee & Duflo, 2007). The specialization necessary to develop a larger scale micro-enterprise (and the exposure that comes from it) is not what the poor are looking for. Research in Zambia found that as much as half of the enterprise owners saw their business as a survival effort and did not show any entrepreneurial traits, while another 30% saw their micro-enterprise as a temporary activity while a market gap existed or until they could find (waged) work (Phillips & Bhatia-Panthaki, 2007).

A second characteristic that differentiates survival and growth-oriented entrepreneurs is that the former barely have an income to satisfy the needs of the household, let alone reinvest profits in the business. Estimating earnings is evidently a very difficult task, but was attempted in Kenya by the Gemini team (Daniels, 2001). For two thirds of the firms surveyed, earnings were even below the minimum subsistence wage set by the government for unskilled workers, which is considered too low for a family to satisfy its basic needs. Entrepreneurs kept their business open, nevertheless, because half of them had this business as a complementary source of income that provided less than half of the household needs.

There is also a gender dimension distinguishing survival and growth-oriented enterprises. Most of the survival-oriented entrepreneurs are women struggling to balance their productive and reproductive roles (Karim, 2001; Marcucci, 2001). They can hardly work on the expansion of their enterprises when they become pregnant again or are overworked in the household, uneducated and unexposed to markets. In addition, they have limited access to social support networks with economic relevance. 'Poor women prefer to expand only to the limits of their own labour and management capabilities', so not expanding their business makes a lot of sense to them (Marcucci, 2001, p. 71).

A final characteristic is their differential access to business development services, credit, and social networks. While survival entrepreneurs may qualify for micro-credit schemes, the amounts loaned are normally too small to allow for growth-oriented investment (Zandniapour et al., 2004). These programmes often target only women, use group lending and group-based collateral arrangements that are not appropriate for growth-oriented entrepreneurs (Richardson et al., 2004). In contrast, a study of evaluation papers of Business Development programmes in Africa, Asia, Middle East, North Africa, Latin America, the Caribbean, and transition countries commissioned by USAID revealed that firms with 10 to 45 workers benefited more from access to credit than micro-enterprises with 1 to 9 workers (Zandniapour et al., 2004).

While graduation of micro-enterprises from the very small scale to a medium size rarely happens, it is still possible. But a closer look at these rather exceptional cases actually reinforces the argument that survival and growth-oriented entrepreneurs are different groups, rather than different stages in the trajectory of firms. An early trace of growth potential among micro-enterprises is the hiring of labour. Even if they start very small, growth-oriented entrepreneurs show the capacity to accumulate capital, create jobs already in the initial stages of the firm, and subsequently expand it vigorously (Mead & Liedholm, 1998). It is precisely this capacity for job creation that makes micro-enterprises worth supporting, but it is also a characteristic that only few of them exhibit (Mead, 1994, p. 1884). In contrast, survival entrepreneurs with surplus revenues typically prefer to create new enterprises instead of enlarging the existing ones (Afenyadu et al., 1999; Nichter & Goldmark, 2009) because they give priority to risk diversification rather than growth (Richardson et al., 2004). This horizontal or lateral growth trajectory (multiplication of enterprises) is seen as a

weaker strategy than a vertical one (expansion of a core business) (Olomi et al., 2001). The salience of micro-credit programmes seems to have exacerbated the preference for horizontal growth due to the ceilings imposed by micro-finance institutions and 'savings groups'.

We fully agree with Mead and Liedholm (1998) that identifying and supporting growth-oriented entrepreneurs is worth the effort. However, given the huge number of survival entrepreneurs and the fact that the overwhelming majority of them will remain in this group and use their business as a buffer more than a means for upward mobility, we feel that more dedicated attention should be given to specific policies for this group. In short, survival entrepreneurs are a different target group, requiring different interventions based on a different logic.

6 Conclusions: Aligning Entrepreneurs and Policies

Existing policies for micro-entrepreneurs are impregnated by a focus on the few of them who have growth potential. This is partly because of a conscious choice to concentrate on those entrepreneurs 'who can make a difference' and provide 'role models' (see e.g. Rogerson, 2001), and partly because almost all enterprise development professionals use characteristics of growth-oriented businesses as outcome indicators for their interventions. Such behaviour leads to a bias in the 'income and employment generation' dimension of poverty alleviation programmes, in which the 'entrepreneurship and business development' component takes on a life of its own. Instead, such programmes should aim at providing a modest but crucial contribution to poverty alleviation interventions by bolstering the role of survival businesses as a buffer against slipping deeper into poverty.

This requires restating a well-established dichotomy within (chronic) poverty reduction schemes between promotion approaches (which aim to increase incomes, productivity, or employment prospects of poor people) and protection approaches (which aim to reduce the vulnerability of the poor) (Matin & Hulme, 2003, p. 647). Just like broader poverty reduction programmes, effective survival entrepreneur policies require both a promotional and a protectional component. Numerous critical studies in the tradition of Portes et al. (1989) and de Soto (2000) have not shattered widespread optimism about the informal economy as an engine of growth.

This optimism has led to a prevalence of promotional strategies and a neglect of protective components, leaving the neediest enterprises without appropriate support.

The enterprise-development intervention practice focuses squarely on promotion and uses an implicit growth-oriented lens. The protection component, a cornerstone of a more critical strand in the informal sector literature (Cook et al., 2008, Lund, 2008), has in recent years entered also the mainstream informal sector policy debate (e.g. see World Bank, 2007). However, the promotion and protection components are not integrated: The promotion interventions continue to start from the mental image of growth-oriented entrepreneurs, while separate protection interventions try to address the needs of informal sector workers and 'self-employed' survivalists.

Enterprise development policies, including those explicitly targeting the smallest types of enterprises and the more informal and poorer entrepreneurs, are implemented at three levels. The least tangible are macro-level policies meant to promote a more enabling environment which is expected to 'unleash' the available entrepreneurial talent among the poor (UNDP, 2004). In its more advanced versions, this would go beyond simply 'levelling the playing field' by including specific measures to counteract existing systematic discrimination against smallness and informality. Once these institutional hurdles have been demolished, survival entrepreneurs are expected to see opportunities for successful businesses and will want to specialize and invest (de Soto, 2000). Such policies are based on the premise that most or at least many survival entrepreneurs are potential winners who simply need a 'break'—like access to microcredit—to lift themselves out of poverty. However, as argued in earlier sections this is not a realistic premise for the overwhelming majority of survival entrepreneurs.

The second set of policies focuses on the meso level of analysis: value chains, clustered economic activities, or local economic development initiatives. These policies go by fashionable labels like 'pro-poor value chain upgrading' (Altenburg, 2007) or 'pro-poor local economic development' (Rogerson, 2006). Especially the value chain and cluster-level policies start from the idea of an 'engine of growth', either larger firms in a value chain or sectorally and spatially clustered groups of firms at local level which generate economic dynamism. The pro-poor dimension is about improving ways to connect smaller and weaker businesses to these 'engines of growth'. Recent policy practice has made significant progress in trying to adapt a value chain-and-upgrading approach to poorer producers (e,g,

see the 'From behind the veil' project at www.bds.knowledge.org), as part of a shift from Business Development Services (BDS) to a 'Markets for the poor' (M4P) approach. They aim to give due weight to the different rationale of survival entrepreneurs by 'incorporating a livelihood security perspective' based on diversification instead of specialization (Miehlbrandt et al., 2005, p. 65). Nevertheless, the M4P approach maintains a basic accumulation stance. The adjustment is to go slower and in smaller steps, providing survival entrepreneurs more time and resources. The goal becomes a more gradual graduation, but graduation from survival to growth-oriented levels of specialization remains the ultimate benchmark. While this may work for some groups and in some situations, it is not a likely generic scenario (Meyer-Stamer, 2006).

The local economic development approach offers more space to accommodate the separate logic of survival entrepreneurs when distinguishing between business development (stimulating the local 'engines of growth') and community economic development, which is 'to facilitate household diversification of economic activity' (Helmsing 2003, p. 69). Nevertheless, LED implementation practice suffers from similar problems as identified by Cotter (1996): Practitioners are divided between those who prioritize economic benefits from the engines of growth and those who focus on issues of redistribution (poverty alleviation, inequality, and exclusion), and too often these groups of practitioners do not systematically cooperate. As a result, the economic dimension of an intervention strategy for survival entrepreneurs has not materialized.

The third and final level of intervention is the micro level of supporting individual entrepreneurs and/or businesses. Much of this type of policies is inspired by the 'missing ingredient' assumption: By providing this ingredient entrepreneurs can start climbing the ladder towards graduation.[4] The two main types of missing ingredients are financial services—including (micro-)credit—and BDS, and these are offered to entrepreneurs in various combinations. The credit-poverty relationship is well-documented; from Paul Mosley's and David Hulme's work (1998) onwards it has been clear that mainstream micro-credit programmes are more suitable to less poor clients. Entrepreneurs with growth potential are better able to use the credit in the assumed fashion, to start and develop a business on which they can increasingly rely for their improving and specialising livelihood.

[4] From our personal experience in visiting and evaluating enterprise development programmes, we found the 'ladder' metaphor to be one of the most popular among project staff.

At the same time, poorer clients use loans more often for coping with crises, are less able to repay along pre-determined schedules, and are more likely to be excluded from group lending schemes. More innovative schemes for poorer clients connect to the rationale of survival entrepreneurs by offering very small and flexible loan facilities at individual level without ordaining clients to use loans for income-generating purposes.

The conventional wisdom in delivering BDS has evolved with a very strong focus on financially sustainable, business-like, and demand-led interventions where entrepreneurs need to prove their anxiousness to absorb a service by paying (almost) a market rate (Donor Committee for Enterprise Development, 2001). While many of these principles make sense for SME programmes, they also effectively prevent access to these services for survival entrepreneurs. As most of them are tailored to the needs of growth-oriented entrepreneurs anyway, this might not be a real issue if and when a parallel system of BDS for survival entrepreneurs would exist. However, while such interventions are attempted in bits and pieces in a wide variety of projects (e.g. see van de Bogaert, 1992), they do not constitute a recognizable set with a common identity or corresponding research tradition.

In all, the existing enterprise and entrepreneurship development policies provide inexcusably few handles for supporting the bulk of survival entrepreneurs, that is, those who are not on the way towards possible graduation. It is as if the elephant in the room has fallen between two stools and become invisible.

A better consideration of the differences between groups of entrepreneurs could have several other ingredients. All policies for supporting survival entrepreneurs have to recognize the 'destructive uncertainty' (Wood, 2003) under which they operate, and primarily aim at increasing security in several dimensions. Bureaucratic bottlenecks, harassment, and corruption affect all enterprises, but it is the smallest and least formal ones that suffer most. Home-based firms in illegal settlements are hampered by deficient infrastructure, inaccessibility for outside customers, and an imminent threat of demolition; street businesses are subject to bribe extortion, temporary or permanent eviction, and confiscation of goods. There is ample evidence that security of tenure leads to increased investment, particularly in immobile and productive assets, and reduces overcrowding in petty trade and services (Berner, 2001; Ghafur, 2001).

Survival entrepreneurs—not surprisingly as most if not all of them are poor—benefit disproportionally from improvements in basic services and

infrastructure. Improving water provision, sanitation and solid waste management and reducing environmental hazards and exposure to disaster diminish the health and time burden that especially poor women have to shoulder; availability of electricity and transport connections mitigates the severe competitive disadvantage that survival entrepreneurs have to cope with; and providing accessible health and education services enhances human capital and contributes to upward mobility for the next generation.

Specific policies to assist survival entrepreneurs to cope better with their insecurities through supporting one or more of their present businesses can only be a modest yet important element in a comprehensive poverty alleviation strategy. Survival entrepreneurs are in no position to pay market rates for credit and BDS, so cost-recovery is in most cases unfeasible. Support for survival businesses cannot be seen as a temporary 'kick-start' intervention which, by providing one or more missing ingredients, set in motion a self-sustaining process of improvement. As a qualitative change in the nature of the business is highly unlikely, many people will be able to better cope with poverty only as long as they receive assistance.

In this chapter we have pulled together earlier attempts to identify a categorical difference among small entrepreneurs and have tried to extend and solidify this basic distinction between survival and growth-oriented entrepreneurs. Secondly, we went beyond the previous literature and tried to systematically develop the policy implications for supporting survival entrepreneurs. We have argued that the present policy practice which claims to target survival entrepreneurs is imbued with pushing entrepreneurs onto the ladder of graduation towards more specialization and growth-orientation. Much further work is necessary, as only through careful research, experimentation, and evaluation can we hope to develop the capacity for 'helping a large number of poor people become a little less poor' (Mead & Liedholm 1998, p. 70)—not so modest a goal after all.

References

Afenyadu, D., King, K., McGrath, S., Oketch, H., Rogerson, C., & Visser, K. (1999). *Learning to Compete: Education, Training and Enterprise in Ghana, Kenya and South Africa*. Education Research Paper No. 42, DFID.

Altenburg, T. (2007). *Donor Approaches to Supporting Pro-Poor Value Chains.* Donor Committee for Enterprise Development. http://www.enterprise-development.org/page/library-item?id=386

Banerjee, A., & Duflo, E. (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives, 21*(1), 141–167.

Berner, E. (1997). *Defending a Place in the City: Localities and the Struggle for Urban Land in Manila*. Ateneo de Manila University Press.

Berner, E. (2001). Learning from Informal Markets: Innovative Approaches to Land and Housing Provision. *Development in Practice, 11*(2 and 3), 292–307.

Billing, K., & Downing, J. (2003). Crossfire: Most BDS Programmes are Wasting Their Time Attempting to Make Survivalist Enterprises Grow. *Small Enterprise Development, 14*(3), 4–7.

Binks, M., & Vale, P. (1994). *Entrepreneurship and Economic Change*. McGraw Hill.

Cook, S., Heintz, J., & Kabeer, N. (2008). Economic Growth, Social Protection and 'Real' Labour Markets: Linking Theory and Policy. *IDS Bulletin, 39*(2), 1–10.

Cotter, J. (1996). Distinguishing Between Poverty Alleviation and Business Growth. *Small Enterprise Development, 7*(2), 49–52.

Daniels, L. (2001). A Guide to Measuring Microenterprise Profits and Net Worth. *Small Enterprise Development, 12*(4), 54–66.

Davis, M. (2006). *Planet of Slums*. Verso.

De Soto, H. (2000). *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. Basic Books.

Donor Committee for Enterprise Development. (2001). *Business Development Services for Small Enterprises: Guiding Principles for Donor Intervention*. http://www.enterprise-development.org/resources/item.asp?resourceid=1

Eigen, J. (1992). Assistance to Women's Businesses: Evaluating the Options. *Small Enterprise Development, 3*(4), 4–14.

Fajnzylber, P., Maloney, W., & Rojas, G. M. (2006). Microenterprise Dynamics in Developing Countries: How Similar Are They to Those in the Industrialized World? Evidence from Mexico. *The World Bank Economic Review, 20*(3), 389–419.

Friedmann, J., & Sullivan, F. (1974). The Absorption of Labour in the Urban Economy: The Case of Developing Countries. *Economic Development and Cultural Change, 22*(2), 385–413.

Ghafur, S. (2001). Beyond Homemaking: The Role of Slum Improvement in Home-Based Income Generation in Bangladesh. *Third World Planning Review, 23*(1), 111–135.

Gómez, G. M. (2008). *Do Micro-enterprises Promote Equity or Growth?* Research Unit Report. Woord en Daad. http://www.woordendaad.nl

Grosh, B., & Somolokae, G. (1996). Mighty Oaks from Little Acorns: Can Microenterprise Serve as the Seedbed of Industrialization? *World Development, 24*(12), 1879–1890.

Helmsing, A. H. J. (2003). Local Economic Development: New Generations of Actors, Policies and Instruments for Africa. *Public Administration and Development, 23*(1), 67–76.

House, W. J. (1984). Nairobi's Informal Sector: Dynamic Entrepreneurs or Surplus Labour? *Economic Development and Cultural Change, 32*(2), 277–302.
House, W. J., Ikiara, G. K., & McCormick, D. (1993). Urban Self-Employment in Kenya: Panacea or Viable Strategy? *World Development, 21*(7), 1205–1223.
International Labour Office (ILO). (1972). *Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya*. ILO.
Karim, N. A. (2001). *Jobs, Gender and Small Enterprises in Bangladesh: Factors Affecting Women Entrepreneurs in Small and Cottage Industries in Bangladesh*. IFP/SEED Working Paper No. 14, Series on Women's Entrepreneurship Development and Gender in Enterprises (WEDGE). ILO.
Leys, C. (1973). Interpreting African Underdevelopment: Reflections on the ILO Report on Employment, Income and Poverty in Kenya. *African Affairs*, 419–429.
Liedholm, C., & Mead, D.C. (1987). *Small Scale Industries in Developing Countries: Empirical Evidence and Policy Implications*. International Development Papers No. 9. Michigan: Michigan State University.
Lund, F. (2008). Social Protection and the Labour Market: Towards a Research Agenda. *IDS Bulletin, 39*(2), 87–92.
Marcucci, P. N. (2001). *Jobs, Gender and Small Enterprises in Africa and Asia: Lessons Drawn from Bangladesh, the Philippines, Tunisia and Zimbabwe*. SEED Working Paper No 18, Series on Women's Entrepreneurship Development and Gender in Enterprises (WEDGE). ILO.
Matin, I., & Hulme, D. (2003). Programs for the Poorest: Learning from the IGVGD Program in Bangladesh. *World Development, 31*(3), 647–665.
Mead, D. C. (1994). The Contribution of Small Enterprises to Employment Growth in Southern and Eastern Africa. *World Development, 22*(12), 1881–1894.
Mead, D. C. (1999). MSEs Tackle Both Poverty and Growth (But in Differing Proportions). In K. King & S. McGrath (Eds.), *Enterprise in Africa: Between Poverty and Growth*. Intermediate Technology.
Mead, D. C., & Liedholm, C. (1998). The Dynamics of Micro and Small Enterprises in Developing Countries. *World Development, 26*(1), 61–74.
Meyer-Stamer, J. (2006). Making Market Systems Work? For the Poor? *Small Enterprise Development, 17*(4), 21–32.
Miehlbrandt, A. O., McVay, M., & Tanburn, J. (2005). *From BDS to Making Markets Work for the Poor*. The 2005 Reader, Annual BDS Seminar, Swiss Agency for Development Cooperation, www.bdsknowledge.org
Mosley, P., & Hulme, D. (1998). Microenterprise Finance: Is There a Conflict Between Growth and Poverty Alleviation? *World Development, 26*(5), 783–790.
Nichter, S., & Goldmark, L. (2009). Small Firm Growth in Developing Countries. *World Development, 37*(9), 1453–1464.
Olomi, D., Nilsson, P., & Jaensson, J. E. (2001). *Evolution of Entrepreneurial Motivation: The Transition from Economic Necessity to Entrepreneurship*. Paper

presented at the Babson College-Kauffman Foundation Entrepreneurship Research Conference.

Phillips, C., & Bhatia-Panthaki, S. (2007). Enterprise Development in Zambia: Reflections on the Missing Middle. *Journal of International Development, 19*, 793–804.

Portes, A., Castells, M., & Benton, L. A. (1989). *The Informal Economy: Studies in Advanced and Less Developed Countries*. John Hopkins University Press.

Portes, A., & Hoffman, K. (2003). Latin American Class Structures: Their Compositions and Change During the Neoliberal Era. *Latin American Research Review, 38*(1), 41–82.

Rempel, H., & House, W. J. (1978). *The Kenya Employment Problem: An Analysis of the Modern Sector Labour Market*. Oxford University Press.

Reynolds, P. D., Camp, S. M., Bygrave, W. D., Autio, E., & Hay, M. (2001). The Global Entrepreneurship Monitor. In *2001 Executive Report*. London Business School and Babson College. www.gemconsortium.org

Richardson, P., Howarth, R., & Finnegan, G. (2004). *The Challenges of Growing Small Businesses: Insights from Women Entrepreneurs in Africa*. IFP/SEED Working Paper No. 47, Series on Women's Entrepreneurship Development and Gender Equality (WEDGE). ILO.

Rogerson, C. M. (1996). Urban Poverty and the Informal Economy in South Africa's Economic Heartland. *Environment and Urbanization, 8*(1), 167–179.

Rogerson, C. M. (1997). Globalization or Informalization? African Urban Economies in the 1990s. In C. Rakodi (Ed.), *The Urban Challenge in Africa: Growth and Management of its Large Cities* (pp. 337–370). United Nations University Press.

Rogerson, C. M. (2001). In search of the African miracle: Debates on successful Small Enterprise Development in Africa. *Habitat International, 25*, 115–142.

Rogerson, C. M. (2006). Pro-poor Local Economic Development in South Africa: The Role of Pro-poor tourism. *Local Environment, 11*(1), 37–60.

Scott, J. C. (1976). *The Moral Economy of the Peasant: Rebellion and Subsistence in Southeast Asia*. Yale University Press.

Standing, G. (1977). Urban Workers and Patterns of Employment. In S. Kannappan (Ed.), *Studies in Urban Labour Market Behavior* (pp. 35–48). International Institute for Labour Studies.

Tellegen, N. (1997). *Rural Enterprises in Malawi: Necessity or Opportunity?* Ashgate and African Studies Centre.

Turnham, D., Salome, B., & Schwarz, A. (Eds.). (1990). *The Informal Sector Revisited*. OECD.

UNDP. (2004). *Unleashing Entrepreneurship: Making Business Work for the Poor*. United Nations Development Programme.

van de Bogaert, M. (1992). Group Entrepreneurship: A Way of Banking on the Strength of the Rural Poor. *Small Enterprise Development, 3*(2), 35–41.

Wood, G. (2003). Staying Secure, Staying Poor: The 'Faustian Bargain'. *World Development, 31*(3), 455–471.

World Bank. (2007). *Informality: Exit and Exclusion: Latin American and Caribbean Studies*. World Bank.

Zandniapour, L., Sebstad, J., & Snodgrass, D. (2004). *Review of Impact Assessments of Selected Enterprise Development Projects MicroReport 3, AMAP Business Development Services (BDS), Knowledge and Practice (K&P) and Research Task Order*. USAID and MDA.

The Developmental Potential of Frugal Innovation Among Mobile Money Agents in Kitwe, Zambia

Iva Peša

> In addition to extending financial services to the poor, mobile money is expected to improve productivity, by increasing the efficiency and lowering the cost of transactions, improving security, generating new employment opportunities, and creating a platform on which other businesses can grow. (Donovan, 2012, p. 62)

Mobile money, especially the example of M-PESA in Kenya, has been hailed as a success story of frugal innovation on the African continent (Paunov, 2013; Knorringa et al., 2016). Frugal innovation, a recent innovation manifestation which is said to be disruptive of innovation processes, business models and even entire economies, can be defined as 'the

I. Peša (✉)
Univeristy of Groningen, Groningen, The Netherlands
e-mail: i.pesa@rug.nl

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_18

stripping and/or (re)design of products, systems or services to make them affordable for low-income customers, without sacrificing user value' (Peša, 2014). If mobile money can cut the costs of transactions and expand financial inclusion, boost economic productivity, foster business growth and generate employment, then it is indeed the quintessential example of frugal innovation. Existing studies on frugal innovation and development have been ideologically polarised. On the one hand 'a business view of 'win–win' in which companies can earn profits while simultaneously alleviating poverty' has been proposed, whereas on the other hand 'critics argue that frugal innovation will merely exacerbate capitalist exploitation and inequality' (Knorringa et al., 2016, p. 144). Whether frugal innovation can lift the Bottom of the Pyramid (BoP) out of poverty through consumption and entrepreneurship or whether it will cause informalisation, aggravate socio-economic inequalities and drive a deeper wedge between developing and developed countries has so far remained a predominantly conceptual question (Chataway et al., 2014; Heeks et al., 2014). Yet 'only through detailed empirical studies can the developmental potential of frugal innovation be properly assessed' (Knorringa et al., 2016, p. 150). This chapter presents one empirical case through which to approach such debates. Based on a qualitative survey among 52 mobile money agents and tellers in Kitwe, Zambia, it poses the question whether the frugal innovation of mobile money can become an inclusive innovation.

Mobile money offers a 'faster, cheaper, and more efficient way […] of moving money' (Hughes & Lonie, 2007, p. 81). Mobile money is a frugal innovation as it provides a simple but effective technology for financial transactions, which radically reduces costs and serves a mass market. Because such a large proportion of the African population remains unbanked, the potential contribution of mobile money to financial inclusion and economic development has been widely proclaimed (Donovan, 2012). Mobile money kiosks have been established in remote rural areas, in peri-urban areas and in informal urban settlements, enabling customers to make financial transactions over long distances, for example by sending money to friends and relatives, conducting business transactions or by buying airtime and paying utility bills (Finscope, 2015). Mobile phone access has increased rapidly over the last five years—in Zambia at least 77 per cent of the population has access to a mobile phone—so entry barriers to use mobile money are low, especially as some services merely require identity documentation for transactions (Habeenzu, 2010). Because mobile money in Zambia has only recently started to reach a mass market,

its impact can still be well discerned. In 2015, 14.0 per cent of the Zambian population was using mobile money services, of which there are three providers: Airtel, MTN and Zoona (Finscope, 2015). Whereas network operators and mobile money customers have been well studied (Hughes & Lonie, 2007; Morawczynski, 2009), mobile money agents and tellers have received relatively less attention. The view even exists that 'agents are […] passive distributors of the product or service: a channel rather than an actor' (Foster & Heeks, 2013, p. 298). Yet by spreading mobile money services to a broad customer base, agents are a crucial link between network operators and customers. One fruitful angle from which to study the role of agents is that of entrepreneurship and labour relations. Mobile money providers, Zoona in particular, prominently advertise stories of successful agents and assert that agents can be (re)made into 'entrepreneurs', because the company offers them 'an opportunity to build their own businesses and create wealth and employment by providing mobile payment solutions, financial services, business management tools and support' (https://www.kiva.org/partners/210). Rather than simply accepting such claims, it should be questioned whether mobile money services are able to generate sustainable employment and entrepreneurship. Can the frugal innovation of mobile money become an inclusive innovation for the agents and tellers in Kitwe, Zambia?

Inclusive innovation can be defined as 'the development and implementation of new ideas which aspire to create opportunities that enhance social and economic well-being for disenfranchised members of society' (George et al., 2012, p. 663). Because 'technological progress and innovation' shape and substantially alter 'how labour and capital are deployed within an economy', this impacts on the distribution of income and processes of socio-economic development (Paunov, 2013, p. 9). Whether innovation works to increase or decrease inequality depends on the innovation itself and on the socio-economic setting within which the innovation is introduced. Studying employment generation and entrepreneurship among mobile money agents in Kitwe provides insights into aspects of inclusiveness. Mobile money providers advertise stories of previously disenfranchised individuals who have become successful agents. But mobile money penetration might give rise to new inequalities, among agents and between agents and tellers. This chapter suggests that frugal innovation can simultaneously have divergent developmental outcomes: it can generate employment and foster entrepreneurship, but it can also contribute to informalisation and create new socio-economic inequalities (Dolan &

Rajak, 2016). Empirical studies thus provide an opportunity to reach a more nuanced understanding of the relationship between frugal innovation and development (Knorringa et al., 2016).

This chapter is structured as follows: First, an overview will be given of debates on mobile money, agents and entrepreneurship, highlighting linkages to frugal and inclusive innovation literature. Secondly, the context of mobile money in Kitwe, Zambia, will be sketched, and the methodology of this study will be explained. In the third section, the results of the survey among Kitwe's agents and tellers will be discussed. In the fourth section, these results will be linked to debates on entrepreneurship, labour relations and inclusive innovation. The fifth and the last section concludes this chapter.

1 Mobile Money, Agents and Entrepreneurship: Some Debates

Mobile phones, by facilitating communication and enabling capital formation, have been identified as 'a key innovative technology in support of livelihoods', notably in ICT4D (Information and Communication Technology for Development) debates (Duncombe, 2014, p. 567). Similarly, it has been asserted that mobile money 'will undoubtedly positively impact the course of economic growth in emerging markets', as it reduces the costs of money transfers and thereby induces savings, investments and remittances, bolsters businesses and boosts employment (Gencer, 2010, p. 102). Mobile money transfers are a uniquely African innovation, born out of the practice of transferring mobile phone credit through SMS (Omwansa & Sullivan, 2012). The first mobile money service, M-PESA, introduced in Kenya by Safaricom in 2007, rapidly obtained a mass market. Only five years after its launch, 70 per cent of Kenyan households had access to M-PESA and there were over 35,000 agents countrywide (Jack et al., 2013, p. 356). Because formal financial services, especially banks, remain largely out of reach of rural and poorer urban Africans, mobile money has been hailed as one means of bringing these un(der)served customers within the reach of financial services, 'banking the unbanked' (Aker & Mbiti, 2010). Part of the attraction of mobile money services is that they are relatively convenient and cheap: they allow users to transfer money to others, make deposits and withdrawals, pay bills and purchase mobile phone credit (Morawczynski, 2009) and they have lower transaction costs than alternative services, such as banks or sending money by bus (Jack & Suri, 2011). Consequently, strong assertions about

the 'transformational' potential of mobile money have been advanced, based on claims that mobile money might 'revolutionise' money transfers, financial transactions and even entire economies on the African continent (Morawczynski, 2009; Heeks, 2014).

Whereas some studies endorse the positive contributions of mobile money to economic development (in terms of income generation, consumption and employment), 'questions remain regarding the nature and extent of m-money's effect on the welfare of poor users in developing countries' (Aker & Mbiti, 2010, p. 222). It should therefore be empirically ascertained how mobile money will affect social divisions of gender, wealth and rural–urban residence. Duncombe (2012, p. 369) more cautiously stresses that 'early adoption of m-finance has favoured those already financially included and market-driven solutions for the financially excluded are limited'. Some authors claim that rather than enhancing development and contributing to inclusion, mobile money and its capitalist underpinnings will merely aggravate the exploitation of the poor, by exacerbating societal and global inequalities (Ya'u, 2004; Murphy et al., 2014). Murphy et al. (2014, p. 265) assert that narratives of financial inclusion and promises of economic development through ICTs 'often overlook the interdependencies and power relations that mediate the potential for African societies to empower themselves within the GIE [global information economy], contingent relations that may exclude some economic actors whilst including, connecting and sometimes exploiting others'. Ya'u (2004, p. 12) critically argues that the digital divide 'defined as unequal access to ICTs within and between nations' is expanding rather than closing, because it 'is part of the wider development divide'. Such existing debates about the developmental potential of mobile money have remarkable parallels to previously mentioned polarising narratives of frugal innovation and inclusive development (Chataway et al., 2014). By empirically engaging with the example of mobile money agents and tellers in Kitwe, this study can contribute to debates on whether mobile money (as an example of frugal innovation) has led to inclusive development, or whether it has merely caused informalisation and a deepening of the 'digital divide' (Heeks et al., 2014).

One useful perspective from which to complicate narratives on frugal innovation, mobile money and development is that of labour relations and entrepreneurship (Meagher et al., 2016). Attention should be paid to the dynamics through which 'global currents in products, services, information and ideologies', might reshape terms of employment and induce local

vulnerabilities (Arora & Romijn, 2012, p. 481). Although mobile money can provide employment prospects, as well as access to skills, capital and technology, Meagher et al. (2016, p. 471) have argued that 'global labour linkages have tended to increase rather than reduce problems of vulnerable and unstable working conditions within African countries'. Critically unpacking the narratives of 'entrepreneurship' among Zambian mobile money agents and paying attention to the unequal labour relations between agents and tellers can contribute to debates about mobile money, frugal innovation and inclusive development (Knorringa et al., 2016). Whereas a number of studies have been conducted on network operators, such as Vodafone and Airtel (Hughes & Lonie, 2007), and on the role of users in appropriating mobile financial services (Morawczynski, 2009), mobile money agents and tellers—in particular their labour relations—have been largely overlooked. Duncombe (2014, p. 573) notes that mobile money agents 'play an important role in delivering services', yet there 'is little evidence from the literature concerning their impact on livelihoods'. This is all the more surprising given that agents are crucial intermediaries between network operators and customers. Agents provide access to mobile financial services through their outlets, they facilitate transactions by managing cash and float and they familiarise first-time users with mobile money transactions, which establishes trust and expands the customer base (Interview #12, Zoona agent manager).[1] Although mobile money providers have celebrated the role of agents in generating employment and fostering entrepreneurship, this has been the focus of few empirical studies so far (Foster & Heeks, 2013; Dolan & Rajak, 2016). Can mobile money indeed contribute to employment generation, entrepreneurship and inclusive innovation? Certain agents have managed to start their own businesses, to open multiple outlets and earn high profits. Yet there is a discrepancy between the fortune of agents and tellers. Tellers, who are employed by agents, carry out mobile money transactions and are responsible for customer relations, they make long hours in the kiosks, yet they are poorly paid and they rarely have a formal contract (Interview #2, Zoona manager). Labour conditions among tellers therefore suggest that 'there are clear limitations to the transformational potential of mobile financial services' (Morawczynski, 2009, p. 521).

Airtel, MTN and Zoona have portrayed mobile money agents as entrepreneurs who can contribute to economic growth and poverty reduction by running profitable businesses (Dolan & Rajak, 2016). Zoona, in

[1] A full list of interviews is provided at the end of this chapter.

particular, depicts 'emerging entrepreneurs' as people 'who dream of starting or growing a business and through personal success can effect real change, drive progress and develop themselves and their communities'. By supporting and unlocking the 'untapped potential' of these entrepreneurs, Zoona allows them to 'earn revenue, create jobs and invest in their communities, resulting in both employment and empowerment'. Following the pioneering work of Schumpeter (1934), economic thought has attributed a transformative role to entrepreneurship in effecting economic change. Entrepreneurship has recently started to receive much attention as a strategy to alleviate poverty (Hall et al., 2012; Alvarez & Barney, 2014). Entrepreneurship might be defined as 'the discovery, evaluation, and exploitation of future goods and services' (Eckhardt & Shane, 2003, p. 336). As 'opportunity entrepreneurs', mobile money agents reacted to a new market opportunity. Whilst first movers acted on 'discovery opportunities', created by technological change and the introduction of mobile money services into the market, subsequent mobile money agents have attempted to consolidate 'creation opportunities', which can serve to build new markets, generate sustained competitive advantage and raise profits (Alvarez & Barney, 2014). Some mobile money agents have succeeded in creating new markets, generating high profits (up to K60,000 per month[2]) and consolidating their competitive advantage through business skills. Entrepreneurship can, in this sense, be good for economic growth. Acs (2006, p. 97) asserts that 'entrepreneurs create new businesses, and new businesses in turn create jobs, intensify competition, and may even increase productivity through technological change'. In this connection, 'improvements in information technologies such as telecommunications [and mobile money] may increase the returns to entrepreneurship' (Acs, 2006, p. 100). More critical studies question the relationship between mobile money and entrepreneurship, arguing that the discourse of entrepreneurship might mask, or even aggravate, structural socio-economic inequalities (Murphy et al., 2014). Dolan and Rajak (2016, p. 515) see mobile money as an 'engine of jobless growth, offering contingent opportunity in place of secure employment'. They assert that 'the experience of entrepreneurship may deliver precisely the opposite of what inclusive business implies', resulting in displacement and inequality, as the labour conditions of tellers suggest (Dolan & Rajak,

[2]The exchange rate was 1 Euro to 10,127 Zambian Kwacha on 1 September 2015: http://www.xe.com/currencytables/?from=ZMW&date=2015-09-01

2016, p. 527). It must therefore be questioned what results entrepreneurship has (on an individual and a broader socio-economic level) and what this implies for the developmental potential of mobile money agents and their businesses. Whether mobile money will prove to be an inclusive innovation from the perspective of labour relations and entrepreneurship can only be assessed empirically. Some background on mobile money services in Kitwe, Zambia, and on the methodology of this study, first has to be provided.

2 A Background to Mobile Money in Kitwe, Zambia

Zoona, which started as a payment system for cotton farmers, was the first company to effectively launch mobile financial services in Zambia in 2009. Offering an interoperable network service, Zoona is one of the few mobile financial services on the African continent that has remained independent of telecom providers (Interview #2, Zoona manager). This implies that Zoona relies solely on mobile money transactions for its profits, as it has no other revenue streams. The company initially struggled to reach commercial viability and scale, relying on external grants to make ends meet. In Kitwe, Zambia's second largest city, kiosks only opened in 2011. In 2012, telecom providers Airtel and MTN, also started to offer mobile money services, thereby diversifying the playing field and increasing the customer base (Interview #1, Mobile technology professional). In 2015, the market share of Zoona agents was estimated at 33 per cent, whereas Airtel and MTN both held 27 per cent of the national mobile money market. The main difference between the three providers is that Zoona works without customer accounts—transactions only require valid identity documentation and a personalised PIN. Airtel and MTN, contrastingly, aim to create a 'network effect' through customer accounts—customers cannot send money without an account, higher fees are charged when transacting money to non-users and intensive marketing and promotions are used to attract customers to join the network (Interview #11, MTN agent manager; Donovan, 2012). Since 2012, mobile money transfers in Zambia have experienced steady growth rates, booming from early

2014 onwards.[3] A Finscope (2015) study estimated that 14.0 per cent of the population makes use of mobile money, but the figure might well be much higher by now (Kabala & Seshamani, 2016).

The majority of Zambian adults (61.8 per cent) lack access to formal financial services, bank accounts in particular (Finscope, 2015). Especially among those who are in informal employment, agriculture or trade, cash remains of paramount importance to any kind of financial transaction. Mobile money has proved a convenient service because of its various uses, ranging from sending and receiving money to/from friends and relatives, to business transactions, paying utility bills, buying mobile phone credit or paying school fees (Kabala & Seshamani, 2016). To transfer money individuals deposit cash at an agent's outlet. The agent converts the cash into e-value (float), which can be transferred to a recipient's phone through an SMS message, containing a unique transaction code. To convert the e-value back into cash, the recipient visits an agent of the same provider and withdraws the money (Maurer et al., 2013). Mobile money thus revolves around the exchange of cash and e-value, which is mediated by agents (Jack et al., 2013). This has caused a degree of financial 'formalisation', 'a shift from informal exchange (through personal remittances, for example) to greater formality through the use of agents' (Duncombe, 2012, p. 571). One agent in Kitwe asserted that from a user perspective, the change from interpersonal money transfer in cash to the introduction of intermediaries in the form of agents and virtual money transfer has been the most impactful shift accompanying mobile money services (Interview #6, Zoona agent). By connecting users to providers, establishing trust in the system and managing mobile money outlets, agents and tellers play a pivotal role in mobile money services.

Mobile money outlets in Zambia are part of larger franchises, which profoundly influences labour relations and opportunities for entrepreneurship. Franchising is a contractual arrangement by which the franchisor (Airtel, MTN or Zoona) contracts individual retailers (agents), offering them dealership of products and services (mobile money) and use of the business format and brand name, in exchange for fees and payments (a share of commissions) (Sigué, 2012, p. 168). In this set-up, Airtel, MTN

[3] In August 2014, Zoona reported having 650+ agents, 600,000+ customers (those who transact at least once every 90 days) and a monthly average transaction volume of $25,000,000+.

Table 1 Customer transaction charges per provider (2015)

	Zoona	*MTN*		*Airtel*	
	Send (free withdrawal)	*Send*	*Withdraw*	*Send*	*Withdraw*
K0–K50	K5	K0.25	K2.50	K0.25	K2.50
K51–K150	K10	K0.25	K2.50	K0.50	K2.50
K151–K300	K20	K0.50	K5	K1	K5
K301–K500	K40	K0.80	K10	K1	K10
K501–K1000	K55	K1	K20	K1	K20
K1001–K5000	K100	K2	K30	K3	K30

Agent commissions are similarly tiered, constituting a small percentage of total transaction charges (less than 10 per cent). Agent commissions are not publicly available, so definitive figures are lacking

and Zoona give the right to use their brand name to individual mobile money agents, who function as independent retailers. Agents have a contractual relationship to Airtel, MTN or Zoona and surrender a share of commissions to the company (See Table 1). Despite the rhetoric of entrepreneurship, agents have never been able to single-handedly manage mobile money outlets. In busy outlets, agents could not serve all customers by themselves, but more importantly, many agents were not prepared to sit in a kiosk full time. Soon the informal practice of hiring 'tellers', who attend to customers and conduct all daily transactions and services, was born. Whereas agents have a contractual relationship to the mobile money provider, tellers do not. Instead, they are directly hired by managing agents, which results in long working hours and poor remuneration. Even if mobile money providers initially disapproved of this practice, it proved a convenient way to scale and the franchising structure left providers with few tools to directly intervene in agent businesses (Interview #3, Zoona customer representative). The franchising structure also enables agents to open multiple outlets. Most agents have between one and three mobile money outlets (with two or three tellers per outlet). The biggest mobile money agent in Kitwe has 13 outlets and employs 25 tellers, but this is the exception rather than the rule. The franchising structure and the attendant discrepancy between agents and tellers influence mobile money businesses, labour relations and entrepreneurship in Kitwe.

The city of Kitwe on the Zambian Copperbelt was chosen as a site to engage with frugal and inclusive innovation, labour relations and entrepreneurship among mobile money agents and tellers. Given high unemployment rates, prospects of employment generation and entrepreneurship through mobile money are particularly welcome in Kitwe. Although the number of outlets has increased rapidly from 2014 onwards, there is still much scope for the expansion and scaling of mobile money throughout Kitwe (Interview #12, Zoona agent manager). After having read the secondary literature on mobile money, frugal and inclusive innovation and entrepreneurship, a qualitative survey for mobile money agents and tellers (Airtel, MTN and Zoona) was designed to tackle the questions: Can mobile money generate entrepreneurship among the agents and tellers in Kitwe? If so, what kind of entrepreneurship is this and who is able to reap its benefits? A qualitative survey, with open-ended questions, was employed as this can determine variation and diversity within a population and it can shed light on meanings and experiences (Jansen, 2010). The survey explored certain predefined topics, which were expected to have a bearing on entrepreneurial ability (such as age; educational and employment background; the process of starting the job and setting up the business; capital formation, training and support; and customer relations, transactions and profits) but it also examined attitudes related to entrepreneurship (benefits and difficulties of being an agent or teller, views on the mobile money business and future plans).This range of topics was meant to establish the empirical diversity in entrepreneurial trajectories among agents and tellers in Kitwe. Between July and September 2015, 52 mobile money agents and tellers completed the survey through interviews, which lasted approximately one hour.[4] The survey was selectively sampled, aiming to provide a spread between agents (23) and tellers (29), mobile money providers (Zoona 24; Airtel 12; MTN 16) and location (city centre 22; residential areas 12; informal settlements 18). The survey results are summarised in Table 2. In order to properly relate these results to their socio-economic context, in-depth interviews were held with ten agents and tellers, to ask them about their career trajectories, business ideas and activities in a more detailed manner. All interviews were transcribed and systematically analysed to look for cross-cutting themes and patterns and to identify common variables. Through multiple sources of information, involving probing, replication and triangulation with secondary data, a context

[4] I thank Edna Kabala-Litana for her invaluable research assistance.

Table 2 Survey results differentiated between agents and tellers

	Agents (23/52)	*Tellers (29/52)*
Age	< 35 (23/23)	< 25 (28/29)
Education	Grade 12 (14/23), college/university (9/23)	Grade 12 (28/29)
Job experience	Mobile sales (7), trade (9), cashier (3)	None (25/29)
Business financing	Social capital (8), previous jobs (11), loans (4)	None
Customer relations	Social construction of a marketplace	Importance of good customer relationships
Franchising	Franchisee, contracted by mobile money provider	Hired by agent, no relation to mobile money provider
Independent business owner versus employee	Independent business owner	Employee
Salary/monthly profits	Up to K60,000	K300–K1500

within which to place mobile money, labour relations and entrepreneurship in Kitwe has been sought. This study provides one empirical case through which to assess 'whether frugal innovation will lead to equitable economic growth or whether it will merely sustain existing inequalities' (Knorringa et al., 2016, p. 148).

3 Labour Relations and Entrepreneurship Among Mobile Money Agents and Tellers

The survey results suggest marked differences between agents and tellers in terms of labour relations and entrepreneurship. This section attempts to explain these differences by focusing on educational and employment background of agents and tellers, financing and business operations and the franchising structure of mobile money in Zambia.

Agents and Tellers: Background

What characteristics differentiate agents from tellers and can these differences explain variations in economic performance? The reception and appropriation of a frugal innovation (such as mobile money) is shaped by intra-community inequalities which determine the ability of individuals to seize economic opportunities (Arora & Romijn, 2012, p. 489). Agents and tellers differ from each other in terms of age, educational level and job

experience and this partially explains their divergent abilities to reap the benefits of mobile money.

Although both mobile money agents and tellers are overwhelmingly young, there is an age difference between them. Whereas all surveyed agents were under 35 years of age, most tellers are even younger and belong to the '18–25 years' age group. This fits into trends of youth informalisation throughout Africa, mentioned by Meagher (2016). Mobile money tellers are predominantly recent Grade 12 graduates without previous job experience (Interview #6, Zoona agent). Their young age facilitates informal labour conditions. Due to high rates of youth unemployment in Zambia (24.6 per cent in 2013), tellers are more willing to accept flexibility and they are less likely to object to low wages or long working hours. Agents, contrastingly, are older because they started out as tellers or in other employment some years ago, gaining experience and accumulating capital before setting up their own mobile money businesses.

In terms of level of education, there is not much difference between agents and tellers when they first start employment. Among both agents and tellers, a Grade 12 certificate is the most common level of education (only one teller had dropped out in Grade 10). Yet what does subsequently differentiate them is that upon becoming an agent and saving enough money, some agents go on to pursue college or university degrees. Becoming an agent thus facilitates the pursuit of higher education (in the fields of business, finance, management or public relations), which in turn creates an advantage over other tellers and agents and aids business expansion (Interview #12, Zoona agent manager). Nonetheless, business success among agents does not appear to be causally linked to level of education. Some agents who earn high profits have a Grade 12 certificate and do not aspire to return to school. One such agent claimed to possess 'business sense', gained through previous work experience in trade (Interview #9, Airtel and MTN superdealer). Significantly, tellers do not lack educational aspirations, and most tellers stated to be saving money for college. Due to high levels of youth unemployment and poverty, working as a teller appears to be an attractive strategy to save money for tertiary education. Yet tellers only rarely regard their employment as part of a 'career path'. Although some tellers aspire to enrol in ICT, business or management courses, many envisage a career in a field unrelated to mobile money, such as education or nursing. Whereas mobile money agents use education to enhance their position as agents and further their business objectives, tellers tend to terminate their employment once they return to

school (Interview #4, Zoona agent and teller). Some agents are thus able to strategically use education within their career paths to attain a certain level of 'elite capture' of the benefits of mobile money, something which tellers are far less well positioned to do (Arora & Romijn, 2012).

Agents and tellers differ markedly when it comes to job experience. Whereas for most tellers their current job constitutes their first employment, all agents have previous job experience. Some agents started out as tellers several years ago, or worked as sales representatives for telecom companies (selling phones, registering SIM cards and advertising promotions). These individuals seized the opportunity to become a mobile money agent, exploiting the 'first mover advantage' as opportunity entrepreneurs (Alvarez & Barney, 2014; Interview #6, Zoona agent). Other agents came into the mobile money business from unrelated occupations. One agent had previously worked as a cashier in a major supermarket chain, whereas another agent had started out as a fish trader (Interviews #7 and #8, MTN and Airtel agents). Prior employment experience can, thus, raise capital and foster the necessary skills to set up as a mobile money agent.

Tellers largely lack such experience, skills and capital, coming straight from secondary school. Still, becoming a teller might be part of a career path, resulting in the opening of a mobile money outlet as an agent. There are examples of tellers who have become agents, predominantly early movers (Interview #4, Zoona agent and teller). Kitwe's biggest agent has promoted three tellers to become agents, with financial support from Zoona. Yet due to a lack of capital, this scenario remains rare. Teller salaries (ranging from K300 to K1500 per month) are not sufficient to save enough money to open a mobile money outlet. Zoona is the only company that offers selective loans to open an outlet, but these loans are exclusively advertised among agents. Moreover, agents have to back the financial reliability of a teller in order to make them eligible for a loan (Interview #12, Zoona agent manager). In an example of elite capture, agents might benefit from keeping an informational gap in place to safeguard their own businesses. Nonetheless, most tellers do not even seem to aspire to become agents. Secondary school graduates might become tellers to avoid unemployment, to contribute to household income or to save money for tertiary education, but only few consider mobile money to be a long-term career path. Most tellers would prefer to further their education rather than set up as an agent, especially because the skills obtained as a mobile money teller are not easily transferrable to other sectors (Interview #5, MTN teller). The result is that most tellers are employed on a short-term

basis (working for six months to one year) under poor labour conditions (working long hours for low salaries). This informalisation drives a further wedge between mobile money tellers and agents. Although agents and tellers are initially similar in terms of age, education and job experience, their contingent career paths subsequently cause a gap in capital, experience and information, which enables agents to more easily capture the benefits of mobile money entrepreneurship. Yet not all mobile money agents perform well and in order to differentiate successful from less successful agents it is important to look at factors of financing, business operations and market construction.

Financing, Business Operations and Market Construction

For agents, financing is a key element in setting up and managing their mobile money business. Zoona requires a minimum capital of K1500, whereas MTN and Airtel require K3000 from agents before opening an outlet, prohibitively large sums of money for most 'emerging entrepreneurs' (Interviews #11 and #12, MTN and Zoona agent managers). Agents typically adopt one of three strategies to raise starting capital. First, agents can save money to open a kiosk through previous employment, for example by working as a teller. Second, some agents use social capital to generate financial resources (one agent was financed by her father's maize mill). Third, Zoona provides a selective loan scheme to boost the entrepreneurial ability of aspiring agents, allowing them to set up their business and repay the loan from commissions (Interview #12, Zoona agent manager). Airtel and MTN, however, do not provide any loans to their agents and hold them responsible for managing their own finances. Initial financing directly affects entrepreneurial opportunities, determining whether an agent can open a high-profit retail shop (a grocery, computer shop or mobile phone outlet where mobile money is combined with other sales activities) or a kiosk (a metal or cardboard structure solely meant for mobile financial services). Apart from starting capital, mobile money relies on continued investment to ensure business growth and sustainability.

Mobile money providers frame agents as entrepreneurs and business owners, responsible for the sound financial management of their own outlets. Yet financial management, especially reinvestment in the mobile money business, often proves to be problematic and differentiates successful from failed agents (Interview #9, Airtel and MTN superdealer). Agents who regard mobile money as a financial opportunity which should be

maximally exploited, who hope to make quick profits by immediately pulling out earnings from their outlets, tend to struggle and some go out of business within six months of opening. If agents run out of cash or float (the store of e-money in an agent's account) they are forced to turn down customers. For example, if a customer wants to withdraw K2000 and this amount of money is not available in the kiosk, the agent or teller has to send the customer to another outlet, causing business loss, reputational damage and decreased profits, especially if this proves a recurring issue (Interview #10, MTN agent). Agents who do invest in building their float can, conversely, gradually expand their business, by conducting large transactions on which high profits can be made (see Table 1). The busiest agents in Kitwe concur that mobile money is not about making quick profits, but about investing in your business. Some agents use side businesses to build their float. One agent sells mobile phones and registers SIM cards, another agent sells handbags, hair and lingerie (Interview #9, Airtel and MTN superdealer). Running multiple businesses can raise profits and spread risks, because if an agent's mobile money account runs short it can be balanced with money from other sales. Sound financial management, indispensable for a successful mobile money business, is not a formal skill learnt through school or training, but is rather an entrepreneurial trait which some agents possess more than others (Interview #6, Zoona agent).

Business success also depends on building the social structure of a new market. The market is a thoroughly social space and 'social structure and relationship development' may well lead 'to improved economic productivity and exchange' (McKague et al., 2015, p. 1086). Good customer relations are fundamental to building a market for mobile money and in this sphere agents and tellers play an indispensable role, by welcoming customers, politely responding to their requests and making small talk (Interview #4, Zoona teller and agent). Especially for first-time users it is important to explain how mobile money works step by step, in order to establish trust and to create customer loyalty. In an instrumental view, agents might be seen as mere tools to expand the customer and profit base for network operators, by facilitating scaling and ensuring commercial viability. Yet agents are social actors, 'their knowledge and social infrastructures are important to the maintenance of the bridges to cash; they add something—credibility, reliability—to the mobile channel' (Maurer et al., 2013, p. 65). Foster and Heeks (2013, p. 309) note that 'the interaction between agents and customers is anything but a simple market transaction, being rooted in a common context of resource constraints, uncertainty, and strong community relationships'. Whereas some agents

have relied on existing social networks, others have built a wholly new customer base for mobile money. Kitwe's first agent established herself in the busy Central Business District. By providing good customer care and ensuring reliable transactions, she attracted many loyal customers (Interview #6, Zoona agent). Another agent targeted minibus drivers, convincing them of the usefulness of mobile money and marketing herself as the most reliable agent (Interview #10, MTN agent). Others used existing social networks to attract customers, relying on location, reputation or previous employment. One agent in a low-income area simultaneously runs a fish business and due to his good reputation he quickly managed to build a customer base and scale his mobile money business (Interview #9, Airtel and MTN superdealer). To carve out and consolidate a market, mobile money services have relied on both pre-existing and new social networks.

The location of a mobile money outlet influences business opportunities. The busiest outlets are located in the city centre, around markets and bus stops (Interview #11, MTN agent manager). Depending on the entrepreneurial initiative of an agent, mobile money outlets can be established in low-income outlying areas. The social network and reputation of the agent is especially important in such locations, as low-income customers are particularly cautious of mobile money. Although customers in outlying areas often have less to spend per transaction, the high volume of transactions can still make business profitable (Interview #9, Airtel and MTN superdealer). Mobile money can thus be scaled by entrepreneurial agents who are willing to establish themselves in low-income areas. Even though financing, business and markets are structured by the Zambian economy at large, entrepreneurial agents have the ability to shape these structures and create profitable businesses.

Franchising and Labour Relations

Numerous benefits of franchising are hailed in the literature, including 'job creation, development of SMEs, economic growth, increase of local entrepreneurship, alleviating poverty, and improving managerial capabilities and skills of the labor force' (Sigué, 2012, p. 168). More negatively, franchising and the agent network of mobile money might simply be another way for service providers to reduce costs, scale among new customers and control risks (Mas & Siedek, 2008). There is a delicate balance of (in)dependence in company–agent relationships. Airtel, MTN and Zoona give the right to use their brand name to mobile money agents,

who function as independent retailers. Agents enjoy a large degree of freedom to run their businesses according to their own insights, even if they do have to follow certain company rules and regulations, for example, regarding transaction procedures, commission structures or security regulations (anti-money-laundering rules). In exchange for any support that agents receive (branding, marketing or training), they surrender part of their commissions to the mobile money provider (See Table 1). Agents are held responsible for managing their own finances, without financial support from mobile money providers. Zoona is the only company that assists its agents through selective start-up loans and float balancing (Zoona Cash), which nonetheless come with stringent repayment conditions. The agent remains responsible for business success or failure and poorly performing agents are not backed by any of the three companies (Interviews #2 and #11, Zoona and MTN agent managers). In order to realise benefits from franchising, agents need support from Airtel, MTN and Zoona in the form of training courses, support staff and financial assistance. The franchising relationship, whereby mobile money providers partially divest themselves from agents and hold them financially responsible for making profits through commissions, can have problematic consequences. In the race for commissions, agents have started hiring tellers for minimal wages, without regulated working conditions. The framing of agents as profit-maximising entrepreneurs by mobile money providers has induced informalised labour conditions among tellers (Meagher, 2016).

The franchise model has created a large discrepancy between agents and tellers, as agents are formally contracted by mobile money providers, whereas tellers are informally employed by agents. By hiring tellers without knowing them personally and without monitoring them closely, agents 'consolidate their positions by actively producing others' informality and vulnerability' (Meagher et al., 2016, p. 475). In an attempt to maximise profits, agents condemn tellers to low salaries, insecure and flexible labour conditions, without long-term contracts and without worker rights (Standing, 2014). These informalised labour conditions are expressed among tellers in numerous ways. First, tellers are poorly paid (between K300 and K1500 per month) and have no prospects of wage increases. Second, tellers work long hours (from 08.00 to 17.00 and beyond, including on Saturdays and Sundays), and they are randomly shifted around various outlets. Third, tellers rarely have a contract which states their rights. Finally, teller turnover is high as most tellers leave their job within one year (Interview #2, Zoona manager). Tellers are relatively educated, but they have no outlook to be

promoted, causing frustration and a lack of motivation. The distinction between agents and tellers in terms of labour conditions was clearly reflected in the survey results, as agents invariably classified themselves as independent business owners, whereas tellers saw themselves as employees. Whilst the franchising model enables agents to open several mobile money outlets, which might boost the entrepreneurial abilities and profit margins of individual agents, it also serves to entrench informalised labour conditions among tellers (Interview #6, Zoona agent). Whereas agents focus on expanding their business, most tellers feel disgruntled with their labour conditions and do not view mobile money as a career path. Understanding this difference between agents and tellers is crucial for engaging in debates on inclusive innovation, entrepreneurship and development.

4 Inclusive Innovation, Entrepreneurship and Informalisation

If inclusive innovation revolves around implementing new ideas and creating opportunities that enhance socio-economic well-being (George et al., 2012), then entrepreneurship, which is the discovery and exploitation of opportunities, appears an ideal tool to achieve inclusivity. This study has questioned whether mobile money can contribute to employment generation, entrepreneurship and (economic) development, by focusing on the labour relations of agents and tellers in Kitwe. Although entrepreneurship is hailed as a 'panacea for inclusive growth', it might have 'unanticipated negative outcomes' (Hall et al., 2012, p. 785). Whereas some agents have indeed managed to act as entrepreneurs, tellers face informal labour conditions. This distinction is in part a result of the franchising structure of mobile money, and can in part be attributed to global trends, whereby 'large firms increasingly make use of casual labor and externalize employment as part of corporate strategies of flexible production' (Lindell, 2010). By exploring the different labour positions of agents and tellers, debates on frugal innovation and inclusive development can be advanced.

Entrepreneurs work to seize opportunities and thereby actively change their socio-economic settings, including resource constraints, such as lack of capital, knowledge or skills (Eckhardt & Shane, 2003, p. 339). Some mobile money agents have succeeded in creating new markets, generating high profits (up to K60,000 per month) and consolidating their competitive advantage through business skills. A study of mobile money agents thus offers an insight into processes of market creation, employment generation

and entrepreneurship. Yet not all agents are equally fortunate. Some agents close down their business within months, whilst others struggle to make ends meet due to the precariousness of their commission-based income (Interview #12, Zoona agent manager). Moreover, the franchising relation creates vulnerabilities, connecting agents to network operators in a skewed and dependent manner. Framed as profit-maximising entrepreneurs, agents are burdened with the responsibility of generating commissions, without receiving much formal assistance from Airtel, MTN or Zoona. Neither does the contract between agents and network operators include any social benefits, such as health insurance or pensions. Being a mobile money agent is thus an inherently unstable profession, profitable whilst things are going well but holding no guarantees for the future. In this connection, narratives of entrepreneurship might 'shift attention from the failure (and profound inequities) of the economic structure in which […] targets exercise limited economic agency' (Dolan & Rajak, 2016, p. 527).

When considering the position of tellers, it becomes even more obvious that 'inequality and displacement can run alongside the BoP promise of inclusion' and entrepreneurship (Dolan & Rajak, 2016, p. 518). Tellers suffer from 'informalisation', 'in having weak or non-existent labour contracts, no 'rights', no protection, and low or volatile earnings' (Standing, 2014, p. 973). Mobile money tellers are not able to save sufficient money to advance their education or career, they do not build transferrable skills and they have no job security. This causes low work morale, which affects customer relationships and jeopardises trust in the mobile money system. Mobile money might therefore be 'beneficial in terms of new livelihoods, but problematic in terms of its longer-term development impacts' (Carmody, 2012, p. 8). Whereas the story of some mobile money agents endorses the business view of a 'win–win' scenario through frugal innovation and inclusive development, the story of most tellers paints a much bleaker view of informalisation and inequality (Knorringa et al., 2016). Departing from polarising views on frugal innovation and development, mobile money might create entrepreneurship for some, informalisation for others and give rise to new socio-economic inequalities and shifting power relations.

5 Conclusion

Existing studies on the relationship between frugal innovation and development have been ideologically polarised (Knorringa et al., 2016). Whilst some have stressed frugal innovation's contribution to development,

others have emphasised processes of adverse incorporation and capitalist exploitation. Through an empirical study of mobile money agents and tellers in Kitwe, Zambia, it has been questioned whether the frugal innovation of mobile money has the potential to become an inclusive innovation. It has been suggested that mobile money can, in some cases, spur entrepreneurship and economic growth among agents. At the same time, mobile money has caused informal labour conditions among tellers, who work long hours for minimal payment. The case of mobile money in Kitwe proposes that frugal innovation cannot be analysed in terms of either development or exploitation. Instead, mobile money might give rise to new and partially unpredictable socio-economic inequalities, among agents and between agents and tellers. Processes of inclusion might be selective, simultaneously fostering differentiation and exclusion (Meagher et al., 2016). Despite the rhetoric of entrepreneurship, agents and tellers have to reckon with the structural inequalities of franchising on which they exert limited control. Detailed empirical case studies, thus, provide an opportunity to reach a more nuanced understanding of the relationship between frugal innovation and development.

Although this is a single case, which cannot easily be generalised, it does suggest that frugal innovation studies should pay more attention to the longer-term and ambiguous outcomes which frugal innovations have on livelihoods and development. Rather than talking about inclusion or exclusion, development or exploitation, these processes might better be perceived as 'structurally integrated with one another in the workings of economies' (Phillips, 2011, p. 385). Moreover, the role of policy in frugal innovation should receive more consideration. In this case, how might policy manage and mitigate the most unequal outcomes of mobile money labour relations, for instance, by establishing a minimum rate of pay or basic labour conditions? How can policies be designed to encourage entrepreneurship as well as social inclusion in low-income settings (Hall et al., 2012, p. 786)? The case of mobile money agents in Zambia suggests that the franchising relationship, although holding some entrepreneurial potential for agents, should be closely monitored to avoid adverse labour relations among tellers. Moreover, network operators need to support agents and tellers with training courses, support staff and financial assistance in order to maximally benefit from mobile money entrepreneurship. All in all, there remains much scope to empirically assess the developmental potential of frugal innovation and thereby to complicate and nuance existing conceptual debates.

Acknowledgements I wish to thank Edna Kabala-Litana for her invaluable research assistance. The Centre for Frugal Innovation in Africa is thanked for the post-doctoral grant which made this research possible. Earlier versions of this chapter were presented at seminars and workshops in the Hague, Leiden and Eindhoven. I am grateful for the feedback and encouragement received during these events. The two anonymous reviewers helped to sharpen the argument and contributed excellent suggestions.

References

Acs, Z. (2006). How Is Entrepreneurship Good for Economic Growth? *Innovations, 1*(1), 97–107.

Aker, J. C., & Mbiti, I. M. (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives, 24*(3), 207–232.

Alvarez, S. A., & Barney, J. B. (2014). Entrepreneurial Opportunities and Poverty Alleviation. *Entrepreneurship Theory and Practice, 38*(1), 159–184.

Arora, S., & Romijn, H. (2012). The Empty Rhetoric of Poverty Reduction at the Base of the Pyramid. *Organization, 19*(4), 481–505.

Carmody, P. (2012). The Informationalization of Poverty in Africa? Mobile Phones and Economic Structure. *Information Technologies & International Development, 8*(3), 1–17.

Chataway, J., Hanlin, R., & Kaplinsky, R. (2014). Inclusive Innovation: An Architecture for Policy Development. *Innovation and Development, 4*(1), 33–54.

Dolan, C., & Rajak, D. (2016). Remaking Africa's Informal Economies: Youth, Entrepreneurship and the Promise of Inclusion at the Bottom of the Pyramid. *The Journal of Development Studies, 52*(4), 514–529.

Donovan, K. (2012). Mobile Money for Financial Inclusion. *Information and Communications for Development*, 61–73.

Duncombe, R. A. (2012). An Evidence-Based Framework for Assessing the Potential of Mobile Finance in Sub-Saharan Africa. *Journal of Modern African Studies, 50*(3), 369–395.

Duncombe, R. A. (2014). Understanding the Impact of Mobile Phones on Livelihoods in Developing Countries. *Development Policy Review, 32*(5), 567–588.

Eckhardt, J. T., & Shane, S. A. (2003). Opportunities and Entrepreneurship. *Journal of Management, 29*(3), 333–349.

Finscope. (2015). Finscope Zambia. http://www.boz.zm/Publishing/77/77_FSD_Zambia_-Final%20II.pdf

Foster, C., & Heeks, R. (2013). Innovation and Scaling of ICT for the Bottom-of-the-Pyramid. *Journal of Information Technology, 28*, 296–315.

Gencer, M. (2010). The Mobile Money Movement: Catalyst to Jump-Start Emerging Markets. *Innovations, 6*(1), 101–117.

George, G., McGahan, A. M., & Prabhu, J. (2012). Innovation for Inclusive Growth: Towards a Theoretical Framework and a Research Agenda. *Journal of Management Studies, 49*(4), 661–683.

Habeenzu, S. (2010). Zambia ICT Sector Performance Review 2009/2010. http://www.researchictafrica.net/publications/Policy_Paper_Series_Towards_Evidence-based_ICT_Policy_and_Regulation_-_Volume_2/Vol%202%20Paper%2017%20-%20Zambia%20ICT%20Sector%20Performance%20Review%202010.pdf

Hall, J., Matos, S., Sheehan, L., & Silvestre, B. (2012). Entrepreneurship and Innovation at the Base of the Pyramid: A Recipe for Inclusive Growth or Social Exclusion? *Journal of Management Studies, 49*(4), 785–812.

Heeks, R. (2014). ICTs and Poverty Eradication: Comparing Economic, Livelihoods and Capabilities Models. *Development Informatics Working Paper Series*, University of Manchester.

Heeks, R., Foster, C., & Nugroho, Y. (2014). New Models of Inclusive Innovation for Development. *Innovation and Development, 4*(2), 175–185.

Hughes, N., & Lonie, S. (2007). M-PESA: Mobile Money for the 'Unbanked' Turning Cellphones into 24-hour Tellers in Kenya. *Innovations, 2*(1–2), 63–81.

Jack, W., Ray, A., & Suri, T. (2013). Money Management by Households and Firms in Kenya—Transaction Networks: Evidence from Mobile Money in Kenya. *The American Economic Review, 103*(3), 356–361.

Jack, W., & Suri, T. (2011). Mobile Money: The Economics of M-PESA. *NBER Working Paper No. 16721.*

Jansen, H. (2010). The Logic of Qualitative Survey Research and Its Position in the Field of Social Research Methods. *Forum: Qualitative. Social Research, 11*(2).

Kabala, E., & Seshamani, V. (2016). Mobile Technology and Poverty Reduction in Zambia: A SWOT Analysis. *IOSR Journal of Economics and Finance, 7*(3), 61–74.

Knorringa, P., Peša, I., Leliveld, A., & van Beers, C. (2016). Frugal Innovation and Development: Aides or Adversaries? *European Journal of Development Research, 28*(2), 143–153.

Lindell, I. (2010). Between Exit and Voice: Informality and the Spaces of Popular Agency. *African Studies Quarterly, 11*(2–3), 1–11.

Mas, I., & Siedek, H. (2008). *Banking Through Networks of Retail Agents.* http://www.cgap.org/publications/banking-through-networks-retail-agents

Maurer, B., Nelms, T. C., & Rea, S. C. (2013). 'Bridges to Cash': Channeling Agency in Mobile Money. *Journal of the Royal Anthropological Institute, 19*(1), 52–74.

McKague, K., Zietsma, C., & Oliver, C. (2015). Building the Social Structure of a Market. *Organization Studies, 36*(8), 1063–1093.

Meagher, K. (2016). The Scramble for Africans: Demography, Globalisation and Africa's Informal Labour Markets. *The Journal of Development Studies, 52*(4), 483–497.

Meagher, K., Mann, L., & Bolt, M. (2016). Introduction: Global Economic Inclusion and African Workers. *The Journal of Development Studies, 52*(4), 471–482.

Morawczynski, O. (2009). Exploring the Usage and Impact of 'Transformational' Mobile Financial Services: The Case of M-PESA in Kenya. *Journal of Eastern African Studies, 3*(3), 509–525.

Murphy, J. T., Carmody, P., & Surborg, B. (2014). Industrial Transformation or Business as Usual? Information and Communication Technologies and Africa's Place in the Global Information Economy. *Review of African Political Economy, 41*(140), 264–283.

Omwansa, T. K., & Sullivan, N. P. (2012). Money Real Quick: The Story of M-PESA. *Guardian Books.*

Paunov, C. (2013). Innovation and Inclusive Development: A Discussion of the Main Policy Issues. *OECD Science, Technology and Industry Working Papers.*

Peša, I. (2014). Frugal Innovation in Africa: Technology, Entrepreneurship and Development in Innovation Processes. *Working Paper Centre for Frugal Innovation in Africa.*

Phillips, N. (2011). Informality, Global Production Networks and the Dynamics of 'Adverse Incorporation'. *Global Networks, 11*(3), 380–397.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle.* Harvard University Press.

Sigué, S. P. (2012). The Promises of Franchising in Africa: The Need for a Critical Examination. *Journal of African Business, 13*(3), 168–171.

Standing, G. (2014). Understanding the Precariat Through Labour and Work. *Development and Change, 45*(5), 963–980.

Ya'u, Y. (2004). The New Imperialism & Africa in the Global Electronic Village. *Review of African Political Economy, 31*(99), 11–29.

List of Interviews

Ryan Silutongwe, Mobile technology professional, Lusaka, 29 July 2015.
Bob Keating, Zoona manager, Lusaka, 30 July 2015.
Kirsten, Zoona customer care manager, Lusaka, 1 August 2015.
Nana, Zoona teller and agent, Kitwe, 7 August 2015.
David Phiri, MTN teller, Kitwe, 12 August 2015.
Sandra Jere, Zoona agent, Kitwe, 14 and 20 August 2015.
Kelvin, MTN agent, Kitwe, 15 August 2015.
Christopher Ngambi, Airtel agent, Kitwe, 21 August 2015.
Sydney Makofi, Airtel and MTN superdealer, Kitwe, 23 August 2015.
Progress Mwansa, MTN agent, Kitwe, 23 August 2015.
Emily, MTN agent manager, Kitwe, 25 August 2015.
Mbamwabi, Zoona agent manager, Kitwe, 27 August 2015.

Dynamics of Mobile Money Entrepreneurship and Employment in Kitwe, Zambia

Edna Kabala

1 Introduction

Debates on the development and success of mobile money as an inclusive finance tool have emphasised its ability to accord users the convenience to easily make transactions. Principally, in many African countries such as Zambia, mobile money has been renowned for enabling payments and money transfers in a fast, cheaper and more efficient manner (Aron, 2018; Donovan, 2012; Hughes & Lonie, 2007; Pesa, 2018). Mobile phone subscriptions with an average coverage of about 90% of the population in Zambia have facilitated the progression of mobile money payments (ZICTA, 2020). Specifically, mobile money payments grew to an annual

E. Kabala (✉)
Copperbelt University, the Southern African Institute for Policy (SAIPAR), Zambia Institute for Policy Analysis and Research (ZIPAR), Economics Association of Zambia, Kitwe, Zambia

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_19

average of 126% in value, from K2.07 billion in 2015 to K49.45 billion in 2019.[1] In 2019, the number of active mobile money wallets had increased to about 14 million and approximately 4.9 million were actively transacting, representing 34% of the total registered mobile money wallets. Under the current circumstances with COVID-19, digitalisation has offered an opportunity for contactless business continuity and may likely to continue supporting the growth of mobile money in Zambia (https://itweb.africa/content/G98YdMLYRnwqX2PD). This chapter examines changes in mobile money landscape in relation to entrepreneurship and labour relations based on key informant interviews surveyed from 52 agents in Kitwe.

2 Changes in Mobile Money Development and Entrepreneurship in Kitwe

Zambia has been dominated by mobile network supported providers of mobile money services, since 2009. The main players at inception of mobile money service diffusion were Zoona, Airtel and MTN money services. While trends in mobile money transactions increased from 2009 until 2015, Zoona maintained its place as the dominant player in the mobile money ecosystem. Zoona was known for attractive compensation of its enterprising agents relative to Airtel and MTN. On average, highest performing Zoona agents in Kitwe were paid about K60,000 while tellers were paid between K1000–K2000 monthly. In stark contrast, Airtel and MTN paid their agents an average of K5000 while tellers earned between K500–K800 per month (Informant #2, interview conducted on 7th June, 2021). Market leadership driven by consistent float availability and premium customer service advantage also enabled Zoona to charge relatively high transaction fees for at least 5 years. This was despite Zoona only offering one line of service in typical money transfers.

Mobile money uptake in the country however experienced drastic changes in agency subscription, extension of services, entrepreneurship and labour relations. Particularly, beyond 2015, Zoona had a turbulent business performance with more clients in Kitwe now shifting to adoption of cheaper and diversified services from MTN and Airtel. Increase in customer activity in the mobile money market was now dominated by MTN and Airtel. The significant increase in customer activity may be attributed to customer retention tactics employed by agents that offered Airtel and

[1] Values expressed in nominal terms.

MTN money services. Agents and tellers shared the codes and processes on how clients could successfully make transactions through their mobile money account by themselves (Informant #2, interview conducted on 7th June, 2021). Further, Airtel and MTN were now branded as multifaceted and more useful services with savings account options relative to Zoona which was only known for money transfer. Zoona thus experienced a fast drop in customer traffic and was soon substituted by Airtel and MTN mobile money services. By the end of 2015, MTN was now identified as the market leader in Kitwe. Many agents interviewed cited that this could have been largely due to a lot of clients already subscribing to MTN as their preferred telecommunication service. Accordingly, informant #2 had this to note:

> Things started changing after 2015 when MTN advanced their expansion services through their super agents. The super agents aggressively advertised the cheap and wide services MTN was offering. … By the end of 2016, people shifted to the Airtel money and MTN services, but much more to MTN since most people in Kitwe use more of MTN. We had lost our clients and business went down. Zoona then changed commission agreements and we started to get 1% of the fee charged per transaction below K5000 and 50% of the charge for amounts above K5000. Our agent commissions dropped and … by the time Zoona wallet which could offer similar services to MTN and Airtel was being introduced, it was too late. Zoona also tried to reduce the prices of the transactions, but still, customers had now moved on to MTN and Airtel mobile money services. This drastically affected our services and even the employment some of us were able to offer. Now I only have one teller who does not even get as much I used to pay my tellers then. She gets K700 per month. (Informant #2, interview conducted on June 7, 2021)

Undoubtedly, MTN took over mobile money service leadership and Airtel became popular in Kitwe due to affordable pricing, customer education efforts and innovation of new products. This may account for some of the factors that led to Zoona's loss in competitive advantage and subsequent extinction beyond 2015. Zoona agents and tellers moved on to join emerging mobile money services. Particularly, former Zoona agents and tellers opted to take up services of MTN and Airtel mobile money. Additionally, some agents took on extra services from new players on the mobile money scene such as Zamtel Money and those offered by banks like Zambia National Commercial Bank (ZANACO) and First National

Bank (FNB). Even with the changes in the mobile money market, agents still earn much more than their tellers. Questions of decent employment that assures a sustainable livelihood from incomes also linger on among tellers and former tellers. In fact, some tellers have no intentions of staying in mobile money employment because of how little their earnings are even with MTN leading the mobile money market. Respondent #3, clarified as follows:

> I was first employed as a Zoona teller and I would be paid K1000 by my boss. Then after 2015, she decided to close my outlet saying that Zoona was not making money and our customers were lost to MTN. I knew that even if my income was little and I planned to only stay on the job before I could go to college, it meant something to me. … So I moved on to find another job with an MTN agent where I still work because I haven't raised enough money for college. I earn 30 percent of the monthly commission we make. 70 percent goes to my boss. Because there are many kiosks now and the commissions are taxed, it's usually not a lot I take home. On average, I take home about K550-K700 while my boss would have about K1200-K1800. It just depends on monthly business but it's not easy. I don't see myself in this job in the near future, what it pays me is not enough to cater for my needs such as saving for school and some parts of my upkeep with high prices of everything now. (Informant #3, interview conducted on June 25, 2021)

The expansion in mobile money services presents new opportunities in entrepreneurship and labour relations. The recognition of the gap between incomes of agents and tellers is a motivating factor, holding other things equal. For example, some tellers who may have enough capital transition into application for their own agency licenses through MTN and Airtel which have substituted Zoona in Kitwe (Informant #3, interview conducted on 25th June, 2021). This realisation, though debatable, may lead to more spread out benefits of mobile money entrepreneurship and employment. Meanwhile, although there has been an increase in customer adoption, there is an evident drop in liquidity within the agent network that signals a loss of quality of the agent network and a decreased profitability in agency business (https://www.uncdf.org/article/6355/bank-of-zambia-and-uncdf-release-the-2019-state-of-the-digital-financial-services-industry-report). This may have some future implications for the returns to mobile money entrepreneurship and employment in Zambia especially urbanised areas like Kitwe.

3 Conclusion

Interviews conducted with key informants among Zoona, MTN and Airtel Money agents suggest that mobile money generated productive employment and entrepreneurship for most agents between 2009 and 2015. Beyond 2015, sustaining gains from mobile money agency services was much dependent on the market leadership of mobile money providers. The dynamics in the mobile money market in Kitwe demonstrate that firms that can offer innovation as well as a broad line of diverse and affordable services are the ones that survive fierce competition. Even in this case, the gap between agents' earnings and those of tellers is still profound. Implicitly, the economic transformation conveyed through mobile money may necessarily encourage casualisation of labour in this saturated market. The vast inequalities in remuneration suggest that socio-economic inequalities may be sustained in labour relations between agents and tellers in the mobile money ecosystem in Kitwe.

References

Aron, J. (2018). Mobile Money and the Economy: A Review of the Evidence. *The World Bank Research Observer, 33*(2), 135–188. https://academic.oup.com/wbro/article/33/2/135/5127166?login=true

Donovan, K. (2012). Mobile Money for Financial Inclusion. *Information and Communications for development, 61*(1), 61–73. https://web.worldbank.org/archive/website01523/WEB/IMAGES/IC4D_2-4.PDF

Hughes, N., & Lonie, S. (2007). M-PESA: Mobile Money for the 'Unbanked' Turning Cellphones into 24-Hour Tellers in Kenya. *Innovations, 2*(1–2), 63–81. http://www.mitpressjournals.org/doi/pdf/10.1162/itgg.2007.2.1-2.63

Pesa, I. (2018). The Developmental Potential of Frugal Innovation Among Mobile Money Agents in Kitwe, ZambiA. *The European Journal of Development Research, 30*(1), 49–65.

ZICTA. (2020). *ICT Statistics Portal and Database*. Lusaka. https://www.zicta.zm/

PART III

Labour Agency and Resilience

Is the Canary Alive and Kicking? Labour's Voice Under Crises

Pedro Goulart, Gianluca Ferrittu, and Raul Ramos

1 Introduction

From the "Age of revolution: 1789–1848" (Hobsbawn, 1996) to the "Human capital [20th] century" (Goldin, 2001), industrial relations, labour and its agency have been at the core of understanding the changes in our lives. Either by taking sides or choosing careers and competences,

P. Goulart (✉)
CAPP, Institute of Social and Political Sciences, Universidade de Lisboa, Lisbon, Portugal
e-mail: pgoulart@edu.ulisboa.pt

G. Ferrittu
Lisbon School of Economics and Management - ISEG, Universidade de Lisboa, Lisbon, Portugal
e-mail: gferrittu@phd.iseg.ulisboa.pt

R. Ramos
Department of Econometrics, Statistics and Applied Economics, AQR-IREA, University of Barcelona, Barcelona, Spain
e-mail: rramos@ub.edu

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_20

we could collectively and/or individually influence our destiny. Also in the start of the period under analysis, as Visser (2012) reminds us, workers' strikes at Gdansk had a role in the fall of the Eastern European Bloc. Expectations were high for the period post-1991, but how did it fare regarding agency? In the mines, canaries were used to detect early signs of deadly gases as an alarm sign. As long as canaries were alive, workers' lives were safe. In the post-Cold War world, with declining unions and with the increase of what was once called atypical forms of work, is the Canary of labour voice alive and kicking?

Earlier expectations were high by 1991. From a technocratic perspective, the end of the Cold War allowed the reallocation of resources to more productive aims. The increasing investment in human capital throughout the world also promised a more productive and prepared labour force and, accordingly, less replaceable and with higher agency. And achievements were great from 1991 to the present in this regard: global adult literacy rate increased from 75% to 86%; and gross secondary school enrolment increased from 52% to 76% (net enrolment: to 66%). If one could not afford education, getting a loan now and pay later was possible. Citizens and civil society were asked to participate in decision-making, increasing their ownership of policies. This was a time where arguments for the primacy of individual abilities, choices and potential predominated for an enhanced agency and domain over one's future. In sum, individual voice seemed to inevitably have benefited from the rise in schooling.

But very soon the context showed characteristics against or at the very least ambivalent towards labour. Education quality varied, skills mismatches between labour demand and supply abounded, and faster skills obsolescence polarised the importance of human capital (Frankema, 2009; Nieto et al., 2015). Trapped in student loans, intermittent jobs make repayment by young professionals more difficult. In some extreme cases, mass schooling could in fact reduce labour differentiation and make labour undifferentiated. In others, the progress in the rest of the world would leave some to stand out even more: the ten countries with lowest adult literacy rate are African, with a notable exception, Afghanistan. In the political sphere, restriction in civil rights following 9/11 Al-Qaeda attacks had also impact on social movements and general agency of labour, given the new tools sometimes misused for other purposes (Gillham, 2011). Most steady material progresses were in China, exactly in a country where labour agency is more constrained. Finally, the declining union density accelerated in this period, reducing the bargaining power of labour, while capital merged and sectorialised thus increasing its leverage.

2 Collective Voice and Agency

For almost 100 years academics have been discussing the demise of unions. Earlier prophecies on their decline can be traced to 1932: "We may take it as probable that trade unionism is likely to be a declining influence in determining conditions of labor" (George Barnett 1932 AEA presidential address, in Kaufman, 2001). Worker organisation in face of the Great Depression and measures promoted by Roosevelt inverted the trend at the time.

Current literature has anchored much of the discussion around the path breaking "What Do Unions Do" (Freeman & Medoff, 1984), who stress the importance of unions in collective voice and wage setting, even if with potential costs. Since then, however, most countries experienced a declining union density which may have harmed labour power in the labour market arena—see Fig. 1. Only five countries in 55 increased union density, namely Indonesia (lower) to Singapore, Brazil, Hong Kong and Iceland (higher). All others saw a decline, with larger declines (over 25 pp.) in former Eastern Bloc European countries (signalled with a triangle), but also in Israel, China, Australia, Ireland and New Zealand. Of the

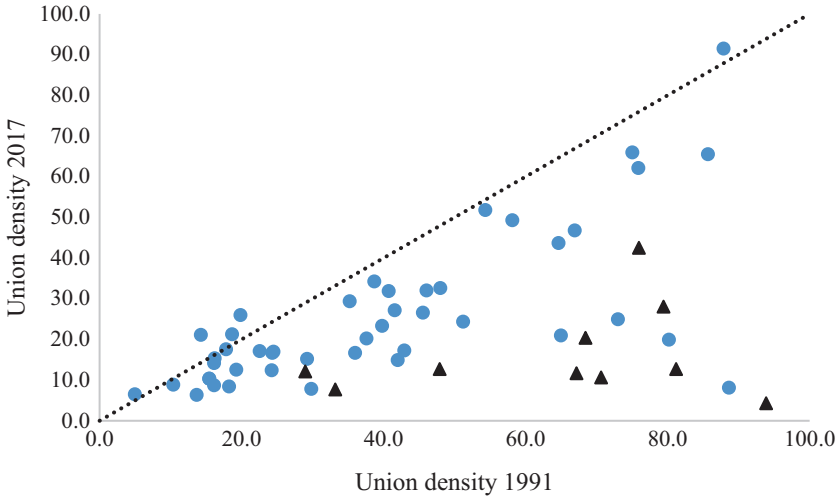

Fig. 1 Change in union density, 1991–2017. (Note: Triangles correspond to former socialist countries. Source: Authors based on Visser [2019])

countries with larger union density in 1991, only Belgium did not face a relevant decline.

Different explanations have been put forward for this decline. Visser (2012) argues that the deindustrialisation and the reduction in collective bargaining paved the way to the decline. Freeman (2005) discusses some potential causes for the decline and emphasises that in the US most workers are still interested in belonging to unions. Rathgeb and Tassinari (2020) argue that the competitive internal devaluations prescribed in Eurozone periphery to respond to the sovereign debt crisis further aggravated the fragility of unions. Trade union strength has been undermined by labour segmentation and, perhaps, the major challenge ahead is bridging the divide between the rights and the voice of insiders and outsiders (Visser, 2019a).[1]

Unions can also be complementary to other organisations and tools in a process of multi-organisation co-production to provide voice to workers. Anner and Liu (2016) find that voice of workers in Vietnam can be channelled by other means other than unions, for example, wildcat strikes, but that happens mostly in environments benefiting from already unionised contexts. Bryson et al. stress the relevance of the context, given that the historical change in union depends crucially on its embeddedness in the economic and social fabric and its need (2011). The Gig economy and its "Just-in-Time Workforce," as put it by De Stefano (2016), present additional challenges that are being addressed by multiple alternatives from the traditional unionisation to alternative fora of collective action (Johnston & Land-Kazlauskas, 2018). The literature also suggests that some negotiating partners are more amicable than others. For example, government seems to be more malleable to wage negotiations with unions than the private sector (Card et al., 2020).

The voice of women and minorities have also been in the spotlight. Also in the academia, there has been increased concern on gender matters in the labour market, see, for example, the well-succeeded nine editions of "The Economics of Women, Men, and Work" (currently, Blau & Winkler, 2021). A review of the literature by Gammage et al. (2016) suggests social norms (different of rules) constrain but can also liberate women. They

[1] This application of the insider/outsider paradigm contrasts to the traditional approach, the latter with considerably more cynicism regarding the role of trade unions, which also needs to consider differences between North America and Europe. See, for example, Lindbeck and Snower (2001).

further document the importance of unions, the public and international organisations in providing voice and reinforcing agency to women.

3 Crises and Resilience

Labour market institutions (including unions) have also been highlighted as one of the factors explaining why some countries (and regions) have better resisted and/or recovered faster than others to different shocks (Blanchard & Wolfers, 2000). The notion of regional and local economic "resilience" has found currency in this literature (Duval & Vogel, 2008). The term "resilience" originated in environmental studies and describes the biological capacity to adapt and thrive under adverse environmental conditions. In particular, a resilient economy is one that retains the capacity to recover quickly from external shocks, even improving its situation from the pre-shock status. In this context, a timely adjustment of wages to the new economic circumstances is a key aspect that requires a flexible wage-setting framework, while at the same time security must be ensured to workers through appropriate labour protection legislation. But impacts on welfare also depend on the capacity of individuals, households and businesses to adapt to the new situation (i.e. micro-resilience) and the different available mechanisms for risk-sharing in any given society.

The Great Recession of 2007/2009 found Euro-area institutions, in contrast to FED policies, more focused on controlling inflation than promoting growth (after an initial zigzag period) leading eventually to "internal" devaluations in Southern European countries. Akerlof's (2001) Nobel Prize address argues that "workers resist nominal wage cuts," an idea already read in Keynes, and managers also avoid these, which would suggest this would be a policy to avoid. The importance of inflation in these moments is that wages can grow above or below average inflation, adjusting for productivity. This inadequate policy would be rectified later on, but only after strong economic, social and political consequences. The crisis also reminded the need to understand the fluctuation of the economy in the short and long run, forcing economists to (re)visit different authors who have tried to explain its origins either by technology (Freeman & Louçã, 2002), microeconomic shocks (Carvalho & Gabaix, 2013) or large firm dynamics (Carvalho & Grassi, 2019).

In contrast to the high expectations and the average growth for all regions, the post-1991 period was an age of turbulent growth (see "Policies for all seasons: The post-Cold War era", Volume II), with

Table 1 Selected economic crises

Year	*Title*	*Coverage*	*Observations*
Early 1990s	Early 1990s Recession	Western Europe, North America, Oceania	"It's the economy, stupid"
Early 1990s	Scandinavia	Finland, Sweden	
Early 1990s	Transition	Most former Eastern Bloc countries	"Shock therapy"
1991		India	
1994		Mexico	
1997	Asian financial crisis	East and Southeast Asia	
1998	Russian financial crisis	Russia	
1999		Argentina, Brazil, Ecuador	
2000–2002	Dot.com bubble	US	
2007–2009	Great Recession	Worldwide	
	Sovereign debt crisis	Europe	
2020–2021	COVID-19	World	

frequent crises disturbing livelihoods and influencing policies and outcomes (see Table 1). In the early 1990s recession, the popularised "It's the economy, stupid" was able to unseat the incumbent George Bush father in favour of Bill Clinton. The Asian crisis in the late 1990s led to costly preventive reserves accumulation. The major shock of the Great Recession eventually pushed policies towards greater equity, but not before the political sustainability of social adjustments implemented after the crisis had been put in question (Goulart & Veiga, 2016). This was reinforced more recently by another life-changing event, the COVID-19 crisis (2020/2021), even if with idiosyncrasies: "[u]nlike the Great Recession, the pandemic has hit women workers harder than men, and disproportionately hurt the job prospects of lower education workers" (Fazzari & Needler, 2021). This stresses the impact on social justice during the recent pandemic (Dosi et al., 2020). Currently, there are worrying signs from China, culminating in Evergrande default. Will the next crisis start in China?

4 In This Section: Labour Agency and Resilience

In this section, "Labour Agency and Resilience," we study these issues by collecting relevant analyses on the evolution of labour agencies and resilience since 1991.

Over the last three decades, global labour agencies had to respond to the complex impacts of megatrends, such as fragmentation of production processes and labour markets integration and segmentation. Despite signs of progress in living standards and in different spheres in human development, noticeably in Asia and particularly China and India, labour markets integration and globalisation of production have distressed the world of labour by neglecting labour markets in developing countries. This has often coincided with high levels of informalisation of the economy, and consequential decline of job quality and security in developing countries. Chapter "Labour and the State in India: Casualisation as Reform" in this book, by Das, contributes to this issue by analysing the challenges faced by the Indian labour market in relation to the increasing informalisation embedded in the Indian neoliberal agenda, which has been implemented since 1991. In the chapter, the author studies the rising informalisation of labour force registered in the last decades in the country, by stressing the counterproductive role of legislative interventions adopted, such as the four Labour Codes, and the increasing feminisation of informal labour. Das (2022) points out the current structural informalisation of the economy, and the lack of strong labour agency in the informal sector, may negatively impact the future developments and achievements of the Indian decent work agenda.

Within the integrated international system of production, negative economic cycles and shocks have often adverse effects on labour markets. These are not confined to the financial sector, but they also impact the lives of people in several ways at the microeconomic level. The adaptive capacity of labour markets to respond and cope with these shocks is different across economic systems and communities, and it is shaped by labour agencies and institutions.

Chapter 22, by Smith, Thomas, Frankenberg, Beegle and Teruel (2022), examines the impacts of the Indonesian 1998 economic crisis on labour markets outcomes. To address this issue, the authors analyse two labour surveys to estimate the long-term employment and wages trends in both rural and urban labour markets, the related impact of the crisis on those, and the role of family strategies in coping with the crisis. In contrast to ILO's emphasis on unemployment, Smith, Thomas, Frankerberg, Beegle and Teruel's findings suggest that the financial crisis in Indonesia lies mainly in the decline of real wages, particularly in the market wage sector. Disaggregating wage cuts, the authors show these disruptive effects have been greater on average in the urban sector, especially at the bottom of the

wage distribution, while cuts in real wages at the top of the distribution are about the same between urban and rural. Smith, Thomas, Frankenberg, Beegle and Teruel's estimates highlight the existence of labour markets flexibilities among the better-skilled workers, and rigidities among the less skilled in times of economic crisis. The authors also suggest that Indonesian households at the top of the income distribution have been resilient to the crisis. At the same time, at the bottom of the distribution, the macroeconomic crisis has had disruptive effects on real incomes. Smith, Thomas, Frankenberg, Beegle and Teruel argue this suggests that the poorest is likely to suffer more the medium- and long-term negative impacts of economic shocks.

Against this background, Chapter 23, by Rafferty, Rees, Sensier and Harding (2022), integrates the analysis on labour markets resilience by examining the pattern of employment and its changes in the North of England during "the long boom" (from 1996 to 2006), and after the Great Recession. In the chapter, the authors implement a logistic regression to explore the regional patterns in the post-crisis period of various labour market indicators, such as involuntary part-time employment and time-related underemployment, and skill under-utilisation. Apart from increased unemployment, Rafferty, Rees, Sensier and Harding (2022) observe unemployment and "broader labour market problems and recessionary impacts" within employment, such as comparatively high levels of labour underutilisation within paid work, especially among younger people, in the North of England. Comparing with the rest of England, the authors also highlight that, in the North, the levels of non-graduate employment among graduates and public sector employment are comparatively high, especially among women. Rafferty, Rees, Sensier and Harding's results suggest that these labour markets disparities between the less industrialised North and the wealthier economic areas of the country (such as the London "super-region"), and among the different economic areas of the North itself, are likely to worsen as a consequence of the crisis. Policy implications for spatially rebalances of the economy are thus discussed in the chapter.

Covariate shocks also impact individuals differently, according to their gender, age and other social characteristics. Gender, in particular, has important implications for the individual capability to cope with labour market risks, which may increase unpaid work situations for women. Social protection programmes and childcare arrangements are thus key factors in

women's ability to work and cope with covariate shocks. Chapter 24 in this book, by Arpino, Pronzato and Tavares (2022), studies this issue by focusing on informal childcare provided by grandparents in Italy and measuring the effect of grandparental childcare on female labour supply. The authors apply an instrumental variable approach, using information regarding whether grandparents are alive to exogenously predict the probability of using grandparental childcare. Arpino, Pronzato and Tavares (2022) find that grandparents' contribution to childcare increases the labour force participation of mothers. This effect seems stronger for less educated mothers, with young children and living in the north and centre of Italy. The authors suggest these results support the idea that grandparental childcare seems to be more used for substituting formal childcare than parental childcare. As such, female labour force participation is likely to increase through formal childcare support policies. In relation to the increasing age of retirement in the labour market, and potential difficulties in providing exhaustive childcare services, Arpino, Pronzato and Tavares (2022) stress the necessity "to frame retirement policies in the larger picture of family and gender policies."

As mentioned before, especially in times of economic crises and uncertainty, age is a fundamental individual characteristic for shocks resilience in the labour markets. Younger people may be more vulnerable than others in these periods, in relation to the policy framework and institutional setting of a given country. Chapter 25, by Eichhorst and Portela (2022), discusses this issue by analysing the impact of economic shocks and labour market institutions' role on employment experiences and perspectives of youth. In the chapter, the authors study the importance of institutions as vocational training systems for youth labour market outcomes in relation to crises' response. By doing so, they focus on the European and developing countries' experiences, discussing the institutional and policy options required to cope with economic shocks and the COVID-19 pandemic crisis. Eichhorst and Portela (2022) argue inclusive labour market institutions may smooth school-to-work transitions, and facilitate the access to better-regulated jobs. At the same time, segmented and exclusive labour markets may negatively impact the transition from the education system, and the characteristics of and access to work for youth. Among the policy options, the authors suggest regulations aimed to improve transitions by increasing vocational and professional training are likely to mitigate the disruptive effects of shocks, and of the unfolding pandemic crisis.

The vulnerability of the middle class has also been in the spotlight and it is a key element to deal with covariate shocks, such as cyclical recessions. It may be important for democracy stability (see Birdsall, 2016), and overall economic development through entrepreneurship and human capital investment (see Acemoglu & Zilibotti, 1997). But despite increasing literature has tried to address middle-class policy issues, there is a lack of what middle class means. In this section (Chapter "A Vulnerability Approach to the Definition of the Middle Class"), López-Calva and Ortiz-Juarez propose an empirical methodology to analyse the middle class based on the notion of vulnerability to poverty, by defining a threshold which ensures that "no lower class or poor households are being identified as middle class," and by establishing a conceptual framework for middle-class analysis over time. In the chapter, the authors propose an absolute-standard definition of the middle class, as those households living with a per capita income of $10–50 a day (PPP, 2005 international $), based on panel data estimates, which considers those households facing a low risk of falling into poverty over time (using a vulnerability dividing line of 10% of probability to fall into poverty). These thresholds are then tested in measuring the size of the middle class in Chile (1992–2009), Mexico (1992–2010) and Peru (1997–2010). The results suggest that in Latin America there is a higher number of middle-class households facing a lower probability of falling into poverty in the late-2000s than in the 1990s than expected. The late-2000s were marked by better economic prospects and social cohesion in the regions, but that would change (see Fig. 1).

The globalisation of labour markets and economic shocks may also be linked with work intensification (see Lu, 2009). This can cause high pressures at work, with ambiguous effects on employees' health and performance (see Hartwig et al., 2020). Workplace resilience and wellbeing programmes have been identified as the main solutions to this novel pressure. Chapter "Trade Unions, Work & Resilience" in this book, by Cotton and Martinez Lucio, contributes to this issue by discussing the landscape and politics of wellbeing and resilience at work, and overviewing the current debates on the implementation of workplace resilience programmes. While the use of these types of interventions has increased since the end of the Cold War, the authors suggest current workplace resilience initiatives, based on the positive psychology model, attempt to pathologise the employees' responses to workplace stress, instead of solving the actual health problems and addressing the triggers. These programmes have thus

been rejected by trade unions. Cotton and Martinez Lucio find the conceptualisation of resilience has become a "contested terrain," and that the dominant strategies about workplace resilience are unlikely to be functioning for coping with the increasing challenges in contemporary workplace regimes, such as work intensification. In this sense, the authors suggest more comprehensive and "radical" research, models and approaches on workplace resilience should be developed.

Natural disasters can cause mortalities and economic and financial shocks, often resulting in significant economic disruptions for the affected regions (see Miao et al., 2018). Disparities and existing policy and socioeconomic settings play a role in this dynamic, shaping the responses and adaptation of agents distressed by the event. Finch, Emrich and Cutter (Chapter "Disaster Disparities and Differential Recovery in New Orleans") conclude this section by studying to what extent the historical sociodemography disparities in New Orleans have shaped the social vulnerability and recovery of local residents from Hurricane Katrina. In the chapter, Finch, Emrich and Cutter apply statistical and spatial approaches for analysing the geographical patterns of the post-disaster response. They contribute by showing empirically that high-social vulnerability and high-flood inundation's combination has a specific spatial dimension, which highlights "the social burdens" of the Katrina disaster. The chapter suggests there are clear spatial disparities in post-disaster recovery and levels of return in New Orleans. Finch, Emrich and Cutter stress the less flooded and less vulnerable areas are recovering faster, while in areas with higher levels of flooding, where the response is slower, differences in post-disaster recovery related to pre-existing social vulnerabilities are detected. Within the range of vulnerabilities, the authors find the middle group is the most hit and the slowest in terms of recovery, while high and low segments of the vulnerable population are recovering faster thanks to private resources and government programmes, respectively.

References

Acemoglu, D., & Zilibotti, F. (1997). Was Prometheus Unbound by Chance? Risk, Diversification and Growth. *Journal of Political Economy, 105*(4), 709–751.

Akerlof, G. (2001, December 8). Behavioral Macroeconomics and Macroeconomics Behavior. Nobel Prize Lecture.

Anner, M., & Liu, X. (2016). Harmonious Unions and Rebellious Workers: A Study of Wildcat Strikes in Vietnam. *ILR Review, 69*(1), 3–28.

Arpino, B., Pronzato, C. D., & Tavares, L. P. (2022). The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Birdsall, N. (2016). Middle-Class Heroes: The Best Guarantee of Good Governance. *Foreign Affairs, 95*(2), 25–32.

Blanchard, O., & Wolfers, J. (2000). The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence. *Economic Journal, 110*(462), C1–C33.

Blau, F., & Winkler, A. (2021). *The Economics of Women, Men, and Work* (9th ed.). Oxford University Press.

Bryson, A., Ebbinghaus, B., & Visser, J. (2011). Introduction: Causes, Consequences and Cures of Union Decline. *European Journal of Industrial Relations, 17*(2), 97–105.

Card, D., Lemieux, T., & Riddell, W. C. (2020). Unions and Wage Inequality: The Roles of Gender, Skill and Public Sector Employment. *Canadian Journal of Economics/Revue canadienne d'économique, 53*(1), 140–173.

Carvalho, V., & Gabaix, X. (2013). The Great Diversification and Its Undoing. *American Economic Review, 103*(5), 1697–1727.

Carvalho, V. M., & Grassi, B. (2019). Large Firm Dynamics and the Business Cycle. *American Economic Review, 109*(4), 1375–1425.

Cotton, E., & Martinez, L. M. (2022). Trade Unions, Work & Resilience. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the post-Cold War era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Das, K. (2022). Labour and the State in India: Casualisation as Reform. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

De Stefano, V. (2016). The Rise of the Just-in-Time Workforce: On-Demand Work, Crowdwork, and Labor Protection in the Gig-Economy. *Comparative Labor Law & Policy Journal, 37*(3), 471–504.

Dosi, G., Fanti, L., & Virgillito, M. E. (2020). Unequal Societies in Usual Times, Unjust Societies in Pandemic Ones. *Journal of Industrial and Business Economics, 47*, 371–389.

Duval, R., & Vogel, L. (2008). Economic Resilience to Shocks: The Role of Structural Policies. *OECD Journal: Economic Studies, 2008*(1), 1–38.

Eichhorst, W., & Portela, A. (2022). Youth Labor Market Integration: The Role of Shocks and Institutions. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Fazzari, S. M., & Needler, E. (2021). US Employment Inequality in the Great Recession and the COVID-19 Pandemic. INET Working Paper No. 154.

Finch, C., Emrich, C. T., & Cutter, S. L. (2022). Disaster Disparities and Differential Recovery in New Orleans. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Frankema, E. (2009). The Expansion of Mass Education in Twentieth Century Latin America: A Global Comparative Perspective. *Revista De Historia Económica / Journal of Iberian and Latin American Economic History, 27*(3), 359–396.

Freeman, C., & Louçã, F. (2002). *As Time Goes By: From the Industrial Revolutions to the Information Revolution*. Oxford University Press.

Freeman, R. B. (2005). What Do Unions Do? The 2004 M-Brane Stringtwister Edition. *Journal of Labor Research, 26*(4), 641–668.

Freeman, R. B., & Medoff, J. L. (1984). *What Do Unions Do*. Basic Books.

Gammage, S., Kabeer, N., & Rodgers, Y. (2016). Voice and Agency: Where Are We Now? *Feminist Economics, 22*(1), 1–29.

Gillham, P. F. (2011). Securitizing America: Strategic Incapacitation and the Policing of Protest Since the 11 September 2001 Terrorist Attacks. *Sociology Compass, 5*, 636–652.

Goldin, C. (2001). The Human-Capital Century And American Leadership: Virtues of the Past. *Journal of Economic History, 61*(2), 263–292.

Goulart, P., & Veiga, F. J. (2016). Portuguese 2015 Legislative Elections: How Economic Voting, the Median Voter and Unemployment Led to 'The Times They Are A'changin'? *Electoral Studies, 43*(3), 197–200.

Hartwig, A., Clarke, S., Johnson, S., & Willis, S. (2020). Workplace Team Resilience: A Systematic Review and Conceptual Development. *Organizational Psychology Review, 10*(3–4), 169–200.

Hobsbawn, E. (1996). *Age of Revolution: 1789–1848*. Vintage.

Johnston, H., & Land-Kazlauskas, C. (2018). Organizing On-Demand Representation, Voice, and Collective Bargaining in the Gig Economy. ILO Working Papers 994981993502676. International Labour Organization, Geneva.

Kaufman, B. E. (2001). The Future of U.S. Private Sector Unionism: Did George Barnett Get It Right After All? *Journal of Labor Research, 22*, 433–458.

Lindbeck, A., & Snower, D. J. (2001). Insiders versus Outsiders. *Journal of Economic Perspectives, 15*(1), 165–188.

López-Calva, L. F., & Ortiz-Juarez, E. (2022). A Vulnerability Approach to the Definition of the Middle Class. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Lu, J. L. (2009). Effect of Work Intensification and Work Extensification on Women's Health in the Globalised Labour Market. *Journal of International Women's Studies, 10*, 111–126.

Miao, Q., Hou, Y., & Abrigo, M. (2018). Measuring the Financial Shocks of Natural Disasters: A Panel Study of U.S States. *National Tax Journal, 71*, 11–44.

Nieto, S., Matano, A., & Ramos, R. (2015). Educational Mismatches in the EU: Immigrants vs Natives. *International Journal of Manpower, 36*(4), 540–561.

Rafferty, A., Rees, J., Sensier, M., & Harding, A. (2022). Growth and Recession: Underemployment and the Labour Market in the North of England. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Rathgeb, P., & Tassinari, A. (2020). How the Eurozone Disempowers Trade Unions: The Political Economy of Competitive Internal Devaluation. *Socio-Economic Review, mwaa021*. https://doi.org/10.1093/ser/mwaa021

Smith, J. P., Thomas, D., Frankenberg, E., Beegle, K., & Teruel, G. (2022). Wages, Employment and Economic Shocks: Evidence from Indonesia. In P. Goulart, R. Ramos, & G. Ferrittu (Eds.), *Global Labour in Distress, the Post-Cold War Era*. Palgrave Macmillan, Springer Nature Switzerland AG.

Visser, J. (2012). The Rise and Fall of Industrial Unionism. *Transfer, 18*, 129–141.

Visser, J. (2019a). Can Unions Revitalize Themselves? *International Journal of Labour Research, 9*(1–2), 17–48.

Visser, J. (2019b, November). *ICTWSS Database*. Version 6.1. Amsterdam: Amsterdam Institute for Advanced Labour Studies (AIAS), University of Amsterdam.

Labour and the State in India: Casualisation as Reform

Keshab Das

1 Introduction: Declining Quality and Availability of Jobs

Since the late 1980s, with over 92 per cent of workers in the informal sector, achieving labour market resilience through reforms has been avidly pursued by the Indian state. A striking contrast is visible in the state's concern for continually 'appeasing' the pampered formal (or organised) sector through numerous fiscal and non-financial incentives while failing to comprehend (hence, to intervene) challenges confronting informal spaces of work. The so-called developmental state has largely focused on the formal industrial and urban sectors and in certain enabling 'modern' information technology services (Raina & Das, 2021: 9). The ongoing economic reforms in India, in place since 1991, have exasperated the employment situation in the manufacturing sector resulting in *jobless* and *job loss* growth as both labour intensity and absolute numbers have fallen. Importantly, in recent years as between 2004–05 and 2018–19, there has been a "massive rise" in both open unemployment and discouraged workers (Mehrotra and Parida, 2021: 306).

K. Das (✉)
Gujarat Institute of Development Research, Ahmedabad, India

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_21

A sustained rise in the contractual jobs even in the formal sector and compromised social security benefits to workers in the unorganised or informal sector have squeezed the quality of employment (Das, 2021: 100). The only sector which has persistently displayed employment elasticity above unity is the construction sector where informality including precarious work predominates. Forced to opt for distress migration from neglected rural areas from chronically underdeveloped states, the large share of labour in urban-industrial India reflects consequences of what the state has not done for labour welfare, creating jobs and local and regional economic regeneration. Labour reforms have cared the least for the informal workers, especially, those inter-state migrants. The existence of at least three prominent national labour Acts (ratified by the Parliament)—Inter-State Migrant Workmen Act, 1979; Building and Other Construction Workers' (Regulation of Employment and Conditions of Service) Act, 1996; and Unorganised Workers' Social Security Act, 2008—has often served little purpose as most informal workers are not registered anywhere and many are unaware of such laws and their rights are hardly addressed. The classic display of the pitiable state of labour laws was seen in May 2020 when millions of migrant workers, having lost their jobs and income as if in a flash, walking back home hungry, sick, carrying kids, bags and without money when a nationwide lockdown was announced abruptly by the central government as a preventive to the spread of COVID-19. While the government failed to come up with even a tentative number of such migrant workers in the country, a rich and painstaking analysis estimated the total (urban and rural) figure of what it terms as the *vulnerable inter-state migrant workers* at a staggering 52 million in 2018 (Srivastava, 2020).

In the absence of appropriate policy regulation, a large supply of labour would, eventually, depress wages and render the labour market flexible whereby the employer (or the owner of the means of production, particularly, capital) is free to *hire and fire*. Contrarily, the shortage of labour whether in number or in terms of a required skill or expertise could exert pressure on the employer who would be obliged to offer higher wages to engage the worker.

2 Rising Informalisation and Feminisation

The prevalence of skewed average earnings between the formal and informal sectors has been a worrying factor as the 'schism' in the labour market has persisted. If, however, the wage adjustments do not take place or do

occur marginally the rigidity of the labour market persists. For instance, it is unlikely that wages (for the same or very similar activities, say, van driving or boiler operating) between the organised and unorganised sectors or, for that matter, wages determined nationally and locally would ever match. Similarly, high unemployment doles or related incentives could act against labour market flexibility.

While overall employment growth has slowed down since 2004–2005, the rising informalisation within the formal sector has been a disturbing phenomenon. The growth of informal workers within formal sector has risen from 6.25 per cent in 2004–2005 to 11.74 per cent in 2018–2019 in the face of an "unprecedented decline in the regular formal workers within the formal sector" (Padhi & Triveni, 2021: 108). This implies that the 'counter-productive' labour laws have made it imperative to undermine formal education and skill of workers who would have to accept jobs on informal terms even within formal enterprises, particularly (Abraham, 2017: 211; Chakraborty, 2015). Another alarming issue has been the sharply divergent income accruing to the capital as compared to those to the workers; for instance, during 1992–2017, the share (of the net value added in organised manufacturing) of wages shows a declining trend while that has risen sharply for profits. This has been the case despite a rise in labour productivity. "This outcome points to the collapse of bargaining power of workers even as the labour market, in line with the capital's longing, has attained a greater degree of flexibility" (Das, 2017: 163–164).

A particularly distressing development, in this process, has been a rising feminisation of the informal work wherein low-end, low-paid and high-drudgery jobs have been available to female workers. Additionally, as shown in Fig. 1, with a declining trend in the female labour force participation rate (LFPR) at least during 1993–2020 concerns have been expressed whether millions of women with limited access to education and skill (especially, those engaged in low-end farm activities) have been pushed to poverty. That the state has a long way to go in designing gender-sensitive labour policies has been argued (Mehrotra & Sinha, 2019).

3 Losing Labour Law Protection

Labour being a concurrent subject (both within national and subnational legislative domains) a huge multiplicity of laws, with overlapping jurisdictions on certain occasions, had emerged over the decades. These laws—presently, 44 national and about 160 subnational/state laws—covered

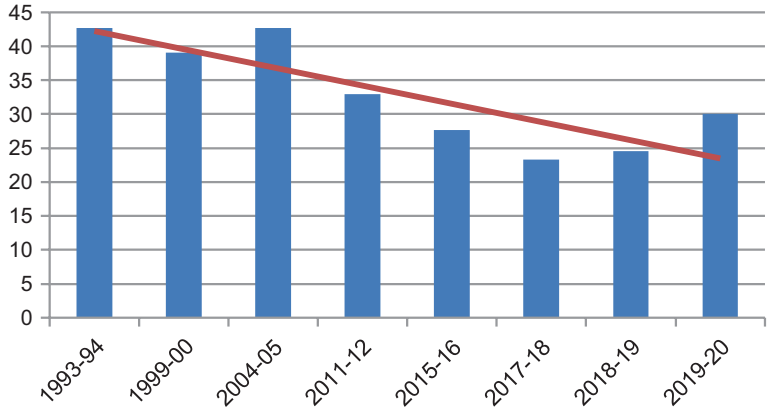

Fig. 1 Female labour force participation rate (per cent) in India, 1993–2020. Sources: For 1993–2016, *Employment Unemployment Survey*, NSSO. For 2017–2020, *Periodic Labour Force Survey*, National Statistical Office, Government of India. Note: 'Usual Status' for the age group 15 years and above

almost all issues concerning right to choose work, eliminating discrimination on social or gender grounds, prohibition of child labour, better conditions of work, social security, wage protection, grievance redressal, right to associate and form trade unions, collective bargaining and participation in management. However, with most informal workers falling outside the purview of these laws and the operational ineffectiveness per se of these, wherever applicable, there were concerns over the salience or relevance of these legislations.

Recently passed in the Parliament the new policy initiative collapses the existing myriad laws into four Labour Codes—Code on Wages (2019), Industrial Relations Code (2020), Occupational Safety, Health and Working Conditions Code (2020) and Code on Social Security (2020). These have virtually nullified several labour rights (including that of collective bargaining) or diluted incentives or benefits by allowing capital to have an upper hand in hiring and firing workers at ease and without impunity. Analysts point to the social dialogue deficit and several loopholes in these hurriedly cleared legislations (Cox & Singhvi, 2020; Sundar, 2021). Inter alia, suspension of labour laws by subnational states would create an additional layer of informalised workers who would be 'pushed out' by their formal employees. "Flexiworkers with limited skills will either be

pushed out of the organised sector or they may be hired only for the disposable nature of their labour. Both scenarios will intensify informality. Workers ousted from the organised sector will crowd into the unorganised sector, thus increasing the supply of labour. In the absence of adequate laws, wages will be driven down" (Sundar & Sapkal, 2020)

The informal workers anyway have little to look forward to from these Codes. The labour agency (especially, the long-established so-called central trade unions) has been a lost cause for them from the beginning as the informal workers are not co-opted as members of these bodies. Even as one looks into the case of informal labour processes within global production networks (GPNs) absence of state regulatory vigilance and non-adherence to labour standards are commonplace, for instance, in subsectors in garments, leather goods, diamond polishing and auto parts sectors. The absence of a labour agency for GPN workers has worsened their bargaining power and wellbeing (Das, 2016, 2020).

4 Concluding Remarks

Deepening of the neoliberal economic reforms has exacerbated conditions of labour who have been doubly disadvantaged by being overwhelmingly informal and without having a responsive agency. The myriad laws as these existed so far have been largely ineffectual in implementation and/or not touched the lives of the informal workers given their nondescript status or 'invisible' contribution. The recently legislated four Labour Codes not only reduce the scope for exercising rights to collective bargaining or social security benefits but have empowered the capital to hire and fire the labour at will. Concerns over ensuring a fair share of wages to workers, participation in management, complying with labour standards and broadening the space for labour activism have been relegated to the background. As the state has been driven by the interests of the capital the decent work agenda would be difficult to advance, especially, as informalisation of work has assumed unprecedented significance in the current political dispensation.

References

Abraham, R. (2017). Informality in the Indian Labour Market: An Analysis of Forms and Determination. *The Indian Journal of Labour Economics, 60*(2), 191–215.

Chakraborty, A. (2015). Reforming Labour Markets in States. *Economic and Political Weekly, 50*(20), 52–57.

Cox, J., & Singhvi, S. (2020). New Labour Codes Dilute Workers Rights and Collective Bargaining. *The Leaflet*, October 1. https://www.theleaflet.in/new-labour-codes-undo-the-law-on-workers-rights-and-collective-bargaining/

Das, K. (2016). Situating Labour in the Global Production Network Debate: As If the "South" Mattered. *The Indian Journal of Labour Economics, 59*(3), 343–362.

Das, K. (2017). Labour Market Resilience in India: Conceptual and Policy Issues. *The Indian Journal of Labour Economics, Special Issue on Labour Market Institutions and Regulations, 60*(2), 155–174.

Das, K. (2020). *Labour Agency and Global Production Networks in India: Intermediaries Old and New*. GIDR Working Paper No. 261, Gujarat Institute of Development Research, Ahmedabad.

Das, K. (2021). Upgrading Technology and Space as Collective Strategy: Creation of Jobs and Market Potential in Gujarat's Ceramic Clusters. In R. Nagaraj (Ed.), *Industrialisation for Employment and Growth in India: Lessons from Small Firm Clusters and Beyond* (pp. 100–128). Cambridge University Press.

Mehrotra, S., & Parida, J. K. (2021). Stalled Structural Change Brings an Employment Crisis in India. *The Indian Journal of Labour Economics, 64*, 281–308. https://doi.org/10.1007/s41027-021-00317-x

Mehrotra, S., & Sinha, S. (2019). *Towards Higher Female Work Participation in India: What Can Be Done?* CSE Working Paper, 2019–02, Centre for Sustainable Employment, Azim Premji University, Bangalore.

Padhi, B., & Triveni, T. (2021). Employment and Informality in the Indian Labour Market: Emerging Trends. *Labour & Development, 28*(1), 93–111.

Raina, R. S., & Das, K. (2021). Whither Rural India? The Developmental State and Labour at the Margins. In P. Swaminathan & U. Rani (Eds.), *State Capital Nexus: Implications for Labour* (Broadsheet on Contemporary Politics) (Vol. 15, pp. 9–12).

Srivastava, R. (2020). Covid-19 and Circular Migration in India. *Review of Agrarian Studies, 10*(1), 164–180. http://ras.org.in/covid_19_and_circular_migration_in_india

Sundar, S. (2021). Labour Law Reforms: After a Tearing Hurry, ad-hocism, Lethargy and Complexities Dominate Processes of Making Regulations—Part I. *The Leaflet*, June 11. https://www.theleaflet.in/labour-law-reforms-after-a-tearing-hurry-ad-hocism-lethargy-and-complexities-dominate-processes-of-making-regulations-part-i/

Sundar, S., & Sapkal, R. S. (2020). Changes to Labour Laws by State Governments Will Lead to Anarchy in the Labour Market. *Economic and Political Weekly, 55*(23), 205. https://www.epw.in/engage/article/changes-labour-laws-state-market-anarchy-labour-market

Wages, Employment, and Economic Shocks: Evidence from Indonesia

James P. Smith, Duncan Thomas, Elizabeth Frankenberg, Kathleen Beegle, and Graciela Teruel

1 Introduction

After over a quarter century of sustained economic growth, Indonesia was struck by a major economic crisis at the end of the twentieth century. The crisis, which worked its way through many of the South East Asian

All correspondence to James P. Smith. We have benefited from the comments of Thomas Bauer, David Lam, Kai Kaiser, Jack Molyneaux and two anonymous referees. We thank the Badan Pusat Statistik, Indonesia, for permission to use their annual labor force surveys. We also thank our colleagues at RAND, Lembaga Demografi, UCLA, and the Population Study Centers in Indonesia who worked with us in the collection of the Indonesia Family Life Survey; we are especially grateful to Bondan Sikoki, Wayan Suriastini, Muda Saputra, and Cecep Sumantri. Financial support from the National Institute of Aging (NIA P01AG08291) and the National Institute of Child Health and Human Development (NICHD 1R01HR40245, 1R01HD33778, 5P50HD12639 and 5P01HD28372) as well as the POLICY Project, World Bank, and World Health Organization is gratefully acknowledged.

J. P. Smith (✉)
Rose Li Associates, Chevy Chase, MD, USA

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_22

"tigers," was presaged by the collapse of the Thai baht in the middle of 1997. During the second half of the year, the Indonesian central bank attempted to stave off pressure on its currency, the rupiah. Interest rates quadrupled and the rupiah depreciated by about 60%. In January 1998, the rupiah collapsed. It lost about 75% of its value in a matter of a few days. The collapse of the rupiah was followed by spiraling prices; annual inflation is estimated to be about 80% for 1998.

The impact of the crisis in Indonesia was not confined to the financial sector. The economy went into a tailspin. Real output in 1998 was estimated to have been about 12% below its 1997 level (a bad year for some Indonesians because of the drought due to El Nino) and economic growth turned positive only in 2000. The crisis was accompanied by dramatic and far-reaching political change.

By any metric, the crisis in Indonesia has been extremely severe—far more severe than in any of its South East Asian neighbors. For example, a year after the crisis began, the Thai baht and Philippine peso stood at about half their pre-crisis values; the Indonesian rupiah had fallen to about one-quarter its 1997 value. A collapse of this magnitude is not unique: the Russian ruble and Ecuadoran sucre were both reduced to one-quarter of their value over twelve months in 1998/1999. Relative to the Russian and Ecuadoran crises, the crisis in Indonesia was different in one key dimension: its timing and virulence were largely unanticipated. For example, in

D. Thomas
Department of Economics, Duke University, Durham, NC, USA
e-mail: dthomas@econ.duke.edu

E. Frankenberg
Department of Sociology, Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA
e-mail: e.frankenberg@unc.edu

K. Beegle
World Bank's Development Research Group, World Bank, Washington, DC, USA
e-mail: kbeegle@worldbank.org

G. Teruel
Universidad Iberoamericana, Mexico City, Mexico
e-mail: graciela.teruel@s2.uia.mx

January 1998, the Government of Indonesia predicted that for the year economic growth would be zero and inflation would be 20%. They were not alone in painting this rosy picture. In July, 1998, the World Bank admitted, "Few predicted the Indonesian crisis, and certainly none its severity.... Indeed, Indonesia seemed better positioned than many other crisis countries on bellwether indicators such as growth, the fiscal stance, the external current account, foreign exchange reserves, and inflation" (World Bank, 1998).

The basic facts of the Indonesian crisis have been widely documented at a macro level (Cameron, 1999). The underlying causes of the crisis, and appropriate macroeconomic remedies, have been the subject of considerable debate (see, e.g., Corsetti et al., 1998a, 1998b; Radelet & Sachs, 1998; Krugman, 1998; IMF, 1999).

Substantially less is known about the magnitude and distribution of the micro-economic impacts on the lives of Indonesians. For example, there is controversy about even basic questions regarding the magnitude of the shock on wages and employment. Some have claimed that the crisis has brought rampant unemployment (ILO, 1998) and called for investments in public works programs (World Bank, 1998). Others have suggested that employment remained remarkably stable through the crisis while real wages collapsed (Frankenberg et al., 1999; Papanek and Handoko, 1999; Feridhanusetyawan, 1999).

Many key issues regarding the effects of economic shocks on labor markets have received little attention in the literature. How do labor markets adjust to a large, abrupt economic decline of the magnitude experienced in Indonesia? Are urban and rural labor markets affected similarly? Do the economic effects differ among the lower, middle, and higher income segments of the labor market? Answering these questions is important not only for documenting the size of the Indonesian crisis and its distributional impact, but also to provide insights into more general questions about the functioning of markets—particularly labor markets—in low-income settings.

While the questions are important, answers are in short supply especially in developing countries where the social science data infrastructure typically has not been in place to monitor the labor market effects of large economic shocks.[1] Fortunately, in this dimension Indonesia represents an

[1] See Fallon and Lucas (1999) for an excellent discussion of the very limited evidence on labor market effects of recent economic crises.

unusual counter-example. First, since the mid-1980s, Indonesia has conducted excellent large-scale, cross-section, nationally representative employment and income surveys, *Survei Amgkatan Kerja Nasional* (SAKERNAS); those surveys are currently conducted annually. Second, during the years proceeding and in the midst of the crisis, a longitudinal survey, the Indonesian Family Life Survey (IFLS), was fielded. The IFLS is an ongoing survey of individuals, families, and communities in Indonesia that offers unusually rich detail about many aspects of the lives of the respondents including their labor market activities. The first wave, IFLS1, was fielded in the second half of 1993, well before the crisis. The second wave, IFLS2, was conducted four years later, in 1997, as the currency crisis was beginning to unfold. Fieldwork for IFLS2 was completed prior to the collapse of the economy and before prices spiraled. In an effort to respond to the scientific and policy needs for quality and timely information on the impact of the crisis, a third round, IFLS2+, was fielded in late 1998, a year after IFLS2. IFLS2+ re-interviewed respondents from a 25% sub-sample of the original enumeration areas. This chapter draws on the sample of almost 2000 IFLS households who were interviewed both in late 1997 and about a year later in 1998. The IFLS data are used in conjunction with 13 years of the SAKERNAS to place the crisis in temporal context and document the nature of the changes that have taken place in Indonesian labor markets in recent years. By laying out these facts, we also lay the groundwork for implementing tests of hypotheses regarding the efficiency and completeness of markets in low-income populations.

The remainder of this chapter is divided into five sections. Section "Introduction" describes some salient dimensions of the social, demographic, and economic changes that have been taking place in Indonesia both over the long-term and during the economic decline of the late 1990s. The second section summarizes the main features of the data used in this research. Section "Data Sources" focusses on labor markets and describes the principal long-term employment and wage trends in urban and rural labor markets as well as how those trends were disrupted by the economic crisis. The next section broadens the analysis to include incomes of the self-employed and highlights the role of dynamics in the labor market; we also examine the role of the family in adjusting to the crisis. The final section highlights the main conclusions.

2 The Indonesian Context

Indonesia, the fourth most populous nation in the world, is an archipelago whose 13,000 islands are home to many different ethnic groups. The country varies a great deal in its urban and rural settlements. For example, population densities range from five people per square kilometer in the province of Irian Jaya to more than 700 people per square kilometer in Yogyakarta. The capital city, Jakarta, has been becoming increasingly important over time as the economic and political center of the country. Jakarta has been a major destination for permanent and circulatory migrants in the last three decades accounting for a quarter of interprovincial in-migrants even though it represents less than 5% of the total population. While not on the same scale as Jakarta, urban centers on other islands also grew in terms of their economic activity and the size of their populations. Still, in spite of the growth of a number of urban centers, Indonesia remains a largely agricultural country. In 1986, a fifth of Indonesians worked in urban areas; this fraction had risen to a third by 1997.

Thirty years ago, Indonesia was one of the poorest countries in the world. Until the recent financial crisis, it enjoyed high economic growth rates and was on the verge of joining the middle-income countries. Prior to the crisis, in terms of size, its economy was comparable to that of Malaysia, the Philippines and Thailand *combined*. On average, GNP per capita in Indonesia grew by 4.5% per annum from the mid-sixties until 1998. However, economic growth was far from uniform across the country with economic heterogeneity having, if anything, increased over time. With the growth of industry and the service sector, the relative importance of agriculture gradually declined (from 55% of total employment in 1986 to 41% by 1997). Over the same years, industrial employment more than doubled from 8% to 19% of the labor force. Not surprisingly, employment in the formal wage sector was expanding, rising from a quarter to a third of all jobs during the same years. This expansion of the formal wage sector, notwithstanding, **it** is important to emphasize that at the start of the economic crisis, most Indonesians workers were either self-employed or unpaid family workers.

Not only has the economy grown rapidly over the last quarter century but there has also been dramatic demographic and social change. Fertility rates have declined substantially and there have been massive investments in human capital. This is reflected, for example, in secondary school enrollment rates which have risen from a mere 6% in 1960 to more than 50% today; primary school enrollment is essentially universal. The investments

are also reflected in the health of the population: at birth, the average Indonesian expects to live to age 61 today which is 50% longer than he or she expected 30 years ago. Thus, the labor force has not only grown dramatically in size but also in quality. The consequences of these improvements are reflected in significant rising real wage levels over the last three decades.

Pride in its past economic achievements and optimism about its future were suddenly challenged by the economic crisis which was accompanied by dramatic shifts in the economic and political landscape of the country. As indicated in Fig. 1, the rupiah came under pressure in the last half of 1997 when the exchange rate began showing signs of weakness. After falling by half from around 2400 per US$ to about 4800 per US$ by December 1997, the rupiah collapsed in January 1998 when, over the course of just a few days, the exchange rate fell by a factor of four to Rp16,000 per US$. Although it soon recovered, by the middle of 1998,

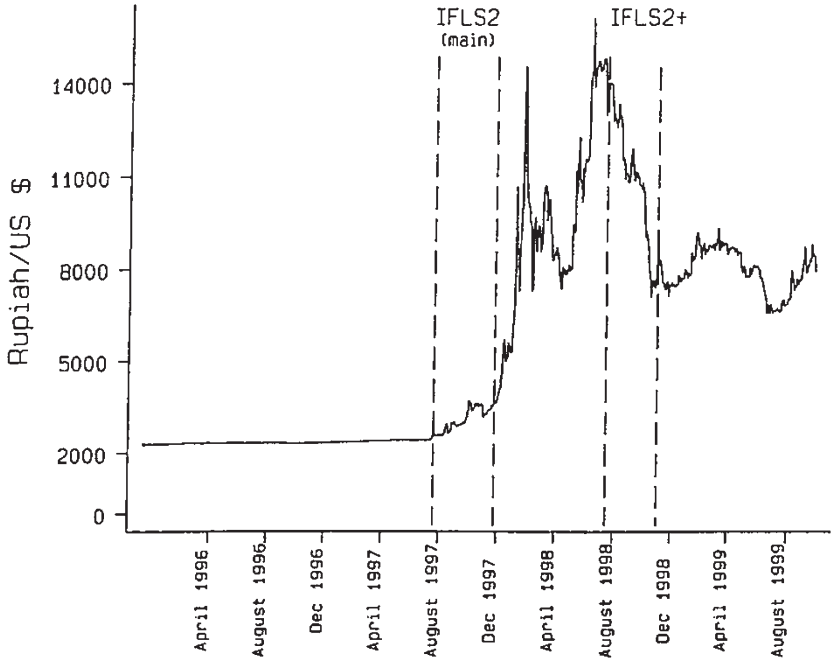

Fig. 1 Timing of IFLS and Indonesian exchange rate

the rupiah had slumped back to the lows of January 1998. Since then, the rupiah has continued to oscillate, albeit at a lower amplitude and frequency. The extremely volatile exchange rate has contributed to considerable uncertainty in financial markets. This is reflected in interest rates which quadruped in August 1997 and were subsequently very volatile. The banking sector fell into disarray and several major banks have been taken over by the Indonesian Bank Restructuring Agency. Turmoil in the financial sector has created havoc with both the confidence of investors and with the availability of credit.

Prices of many commodities spiraled upwards during the first three quarters of 1998. Annual inflation is estimated by the *Badan Pusat Statistik* (BPS), the central statistical bureau, to be about 80% for 1998. In part, this reflects the fact that subsidies were removed on several goods—most notably rice, oil, and some fuels. Part of the inflation can be attributed to the fact that rice accounts for a substantial fraction of the average Indonesian's budget and its price more than doubled. Since the share spent on rice is greatest for the poorest, inflation likely had a bigger impact on the purchasing power of the poorest. Off-setting that effect, however, is the fact that some of the poorest are rice producers and as the price of food rose, so (net) food producers have benefitted from the improvement in their terms of trade. More generally, there is abundant evidence that price increases have not been uniform—either across goods or across space. Inflation estimates produced by BPS are based on prices collected in urban areas; there are reasons to expect prices in rural areas have not moved in lock step with urban prices. Given the magnitudes of the swings in relative prices, it is likely that households will have responded to the price changes by substituting away from goods that have become relatively expensive towards goods whose prices have increased less. These behaviors are not reflected in the BPS inflation estimates (which are based on a basket of commodities calculated in 1996) although evidence from IFLS indicates that, in fact, households throughout the income distribution have altered their spending patterns in response to the relative price changes. Since the BPS province-specific price series is the only series that spans the time period considered in this chapter, for the sake of consistency, incomes are deflated with that series throughout. The issues raised above should be kept in mind when interpreting the evidence (for further discussion, see Thomas et al., 2000b; Levinsohn et al., 1999).

The hypothesis that net food producers may have been partially protected from the effects of the crisis needs to be tempered since a severe

drought immediately preceded the financial crisis and it affected agriculture in many parts of the country—particularly in the east. Country-wide, rice production fell by 4% in 1997 with rice and soybeans being imported. Moreover, unusually severe forest fires raged in parts of Sumatra, Kalimantan, and Sulawesi affecting many aspects of economic life including agriculture and tourism.

Indonesia has also undergone dramatic transformation in the political sector. After thirty years of power, Suharto resigned from the Presidency in May 1998 in the midst of violent demonstrations throughout the country. Multi-party elections in 1999 diminished the influence of the military in political affairs and the granting of independence to East Timor may herald the re-birth of democracy in Indonesia although several other parts of the country are agitating for autonomy and political, ethnic, and religious tensions continue to rock the nation.

Few Indonesians have been untouched by the crisis. For some, the impacts have been devastating, but for others, the crisis has likely brought new opportunities. Exporters, export producers, and food producers may have fared far better than those engaged in the production of services and non-tradeables or those on fixed incomes. Moreover, some individuals may have more quickly embraced new opportunities or more efficiently offset the effects of the economic shocks that they have faced. Given the complexity and multi-faceted nature of the crisis, it is only with sound micro-level empirical evidence that it becomes possible to fully explore the nature and extent of behavioral responses by individuals and families to the crisis and, thereby, characterize with much confidence what the combined impacts of the various facets of the crisis have been and how they have varied across socioeconomic and geographic strata. Moreover, the massive upheavals in the Indonesian economy—and the diversity of their impact—provide an unparalleled opportunity to better understand the dynamics of urban and rural labor markets in low-income settings as well as mechanisms used by families to smooth out the effects of large, unanticipated shocks.

3 Data Sources

This study relies on survey data from two complementary sources: SAKERNAS and IFLS. IFLS is a large-scale integrated socioeconomic longitudinal survey containing extensive information on the lives of respondents, their households and families, as well as on the communities

in which they live. The first wave was conducted in 1993 (IFLS1), with a follow-up in 1997 (IFLS2) and a special follow-up (of a 25% sub-sample) in late 1998 (IFLS2+).

IFLS

The IFLS1 sample is representative of about 83% of the Indonesian population living in 13 of the 27 provinces in the country.[2] Of the households included in the original listing for IFLS1, more than 93% were interviewed yielding a sample of 7224 households. IFLS1 was fielded between August 1993 and January 1994. In each household, representative members (typically the female and male household heads) provided detailed household-level demographic and economic information. In addition to these respondents, several household members were also randomly selected and they were each asked about their own lives. These "main" respondents provided detailed information on a broad array of topics including education; migration; assets and wealth; use of health care and health status; marriage; fertility and contraception. The work history module contains questions on many dimensions of the respondents' current and prior labor market behavior, including income from earnings, hours or work and type of work in both their primary and secondary jobs. Questions are asked about wage employment, self-employed activities, and work in family businesses. The work histories span the 5 years prior to the survey.

The second wave of the IFLS (IFLS2) was fielded between August 1997 and February 1998. Most of the interviews were completed by December 1997; the first two months of 1998 were spent tracking down movers who had not already been found (Frankenberg & Thomas, 2000 describe the survey). Our goal was to recontact all the original IFLS households and re-interview all IFLS1 "main" respondents (as well as all respondents born prior to 1967 independent of whether or not they were "main" respondents); we will refer to these as our "target" respondents. If a "target" respondent had moved, information was sought about the

[2] The sample includes four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five of the Javanese provinces (DKI Jakarta, West Java, Central Java, DI Yogyakarta, and East Java), and four provinces covering the remaining major island groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi). The IFLS1 sampling scheme balanced the costs of surveying the more remote and sparsely populated regions of Indonesia against the benefits of capturing the ethnic and socioeconomic diversity of the country.

person's new location and, if that location was in any of the 13 IFLS provinces, attempts were made to interview the respondent in the new location. Excluding the households in which everyone had died (mostly single-person households), 94% of the IFLS1 households were re-interviewed and individual interviews were completed with 93% of target respondents. In IFLS2, household-level information was usually provided by the household head and spouse and we attempted to complete an individual interview with every household member; over 33,000 individuals were interviewed (representing over 95% of all household members). There are 7600 households in IFLS2; the increase relative to IFLS1 arises because respondents who had split off from the original household and set up their own households were followed.

IFLS2+ was fielded a year after IFLS2, between August and December, 1998. A key goal was to provide insights about the likely immediate effects of Indonesia's economic crisis by collecting high-quality, timely data on who had been affected and on the strategies individuals, households, and communities adopted to mitigate the impact. Because there was neither the time nor resources to mount a survey of the same magnitude as IFLS2 (which took more than two years to plan and test), a scaled down survey was administered, while retaining the intellectual core of IFLS2. In IFLS2+, a 25% subsample of the IFLS enumeration areas (EAs) were selected and all households who had lived in those EAs in 1993, including split-offs followed in 1997, were included in the target sample. The EAs were chosen to be representative of the entire IFLS sample.[3]

Counting all original households in IFLS1 and the split-offs in IFLS2, there are 2066 households in the IFLS2+ target sample. In spite of the turmoil in Indonesia during 1998, the fieldwork was remarkably successful. We relocated and re-interviewed more than 95% of the target households (including 60% of the IFLS1 households that had not been interviewed in IFLS2). The focus in this chapter is on a comparison of behaviors in IFLS2, with behaviors in IFLS2+, about a year later. Taking as our sampling frame those households interviewed in 1997, over 98.5% were re-interviewed in 1998 and 95% of the 1997 respondents completed

[3] The sample was drawn in two stages. To reduce costs, seven of the 13 IFLS provinces were selected: two on Sumatra (North and South Sumatra), three on Java (DKI Jakarta, West and Central Java), West Nusa Tenggara, and South Kalimantan. These provinces span the full spectrum of socio-economic status and economic activity in the fuller IFLS sample. Second, within those provinces, we purposively drew 80 EAs with weighted probabilities to match the IFLS sample as closely as possible.

individual interviews in 1998. These "panel" respondents are the core of the data for our analyses based on the IFLS.[4] Because the same respondents were interviewed twice, once just prior to the collapse of the rupiah in January 1998, and once during the crisis, a comparison of IFLS2 and IFLS2+ allows us to examine the dynamics of the labor market behaviors of individuals and families as the crisis progressed (see Thomas et al., 2000c, for a detailed discussion of attrition).

SAKERNAS

SAKERNAS is the principal Indonesian Labor Force survey administered by the BPS to compile and monitor provincial employment statistics. It has been conducted since 1986. Prior to 1995, the survey was conducted quarterly and annually in August after 1995. The survey covers a nationally representative sample of approximately 40,000 working-age individuals. The survey has been conducted every year except 1995 when the *Survei Penduduk Antar Sensus* (Intercensal demographic survey, SUPAS) was fielded instead. SUPAS covered the same questions on labor force activities as SAKERNAS on a sample that was three times bigger than SAKERNAS and so it is included in our time series of cross-section surveys.

The surveys record demographic information, including age, sex, and marital status, of all household members. For those 10 years of age and older, information is collected on educational attainment and activities during the previous week. For those who were working, questions are asked about days and hours of work as well as occupation and industry of employment. Monthly and weekly earnings are collected for all "employees" or wage earners in the household.

SAKERNAS has several key advantages for our purposes. First, it provides an excellent base to lay out the main facts about labor market changes in Indonesia in the last 15 years. It will be used to cross-validate the main patterns of labor market changes observed in IFLS2/2+. Second, SAKERNAS spans a long enough time period so that the 1998 economic crisis can be given some temporal and economic perspective. Third,

[4] IFLS2 contains over 21,000 individuals who completed the work history module. Of them, 5345 respondents are included in the IFLS2+ target sample based on their EA of residence in 1993. In total 7500 respondents completed the work history module in IFLS2+; of them, 2434 are new respondents (not interviewed in 1997) and 5067 are panel respondents interviewed in both 1997 and 1998.

sample sizes are large so that small changes in outcomes can be measured precisely. The two principal disadvantages of SAKERNAS are that it contains no longitudinal information so that change at the individual or household level cannot be measured directly and it collects no information on incomes earned from self-employment. By drawing on IFLS2/2+, these disadvantages will be addressed.

A potentially important difference between the IFLS and SAKERNAS involves the extent of their geographical representation. While SAKERNAS is nationally representative, IFLS1 is limited to 13 provinces and IFLS2+ was further restricted to seven of those provinces. These restrictions are far less severe than a simple province count would indicate since the most populous provinces were all part of IFLS1, IFLS2, and IFLS2+. Nonetheless, a legitimate question does arise about how comparable the two sets of surveys will be in their depiction on the economic crisis in light of their different geographical coverage.

To address this question, Table 1 lists some key labor market indicators on which we will rely heavily in the rest of this chapter. The table uses SAKERNAS to measure the rate of growth of employment (by sector of work) and hourly wages in the formal (or wage employment) sector between 1986 and 1997 (in the left hand panel) and between 1997 and 1998 (in the right hand panel). The first column in each panel includes all 27 provinces in Indonesia[5]; the second column restricts the SAKERNAS data to the 13 IFLS provinces and the third column to the seven IFLS2+ provinces.

Whether changes in wages or employment rates are considered, the general characterization of growth in Indonesia in the decade preceding the crisis or of the labor market consequences of the crisis is virtually identical no matter whether the complete or more limited set of IFLS2+ provinces are used. While reassuring, this result should not be surprising since over 80% of the Indonesian population is represented in the IFLS sample frame and the additional geographic restrictions necessary for IFLS2+ were explicitly chosen to maintain the representativeness of the sample.

Consider first real wage changes in the formal wage sector.[6] Between 1986 and 1997, mean real wages of males expanded by slightly more than 40%. This increase over more than a decade was almost entirely wiped out

[5] The estimates include East Timor which was part of Indonesia at the time.

[6] Recall self-employment income is not recorded in SAKERNAS. Wages are computed as the sum of monthly cash and in-kind income earned divided by 4.33 times weekly hours. All

Table 1 Labor market changes between 1986, 1997, and 1998 and representativeness of IFLS provinces

Provinces:	*% change between 1986 and 1997*			*% change between 1997 and 1998*		
	All	*IFLS*	*IFLS2+*	*All*	*IFLS*	*IFLS2+*
	(1)	*(2)*	*(3)*	*(1)*	*(2)*	*(3)*
MALES						
Wage	40.3	41.5	42.2	-37.8	-38.1	-36.0
Urban sector	36.2	37.1	37.3	-40.6	-40.9	-38.6
Rural sector	38.2	38.7	39.4	-35.6	-35.9	-33.7
% working	-0.5	-0.7	-0.3	-0.9	-1.1	-1.7
In wage sector	7.1	7.1	6.4	-3	-2.7	-2.5
Self-employed	-3.3	-3.8	-3.5	1.5	1	1.3
Unpaid family	-4.1	-4	-4.6	0.6	0.6	0.6
FEMALES						
Wage	62.1	63.9	67.7	-37.9	-38.3	-39.2
Urban sector	62.5	64.8	68.7	-41.3	-41.8	-43.3
Rural sector	52.6	52.6	52.7	-33.9	-33.8	-32.7
% working	-2.3	-0.9	-1.3	1	1.3	1.0
In wage sector	3.7	3.8	3.8	-0.3	-0.2	-0.2
Self-employed	1.1	0.6	1.2	0.5	0.4	0.5
Unpaid family	-6.4	-5.5	-6.1	0.8	1.1	0.8

Source: SAKERNAS (1986–1998). IFLS provinces are North Sumatra, South Sumatra, West Sumatra, Lampung, Jakarta, West Java, Central Java, East Java, Yogyakarta, Bali, West Nusa Tenggara, South Kalimantan, South Sulawesi. IFLS2+ provinces are North Sumatra, South Sumatra, Jakarta, West Java, Central Java, West Nusa Tenggara, and South Kalimantan. Wage changes calculated for those who earned a wage in the market wage sector only; wages measured in 1997 rupiah using BPS province-specific deflators. All estimates weighted by province weights

in one year during the economic crisis. Long-term wage growth of males was only slightly lower in the urban sector, but the urban sector was hit more severely by the economic shock. Between 1997 and 1998, wages of males declined by about 5% more in urban areas so that real wages were actually lower in 1998 than in 1986. In the formal wage sector, women had actually enjoyed greater wage increases in the years before the crisis. Between 1986 and 1997, wages of females rose by more than 60% thereby reducing the female/male wage gap in paid employment by more than

wages are in terms of 1997 rupiah deflated using province-specific consumer price indices published by BPS.

20%. The crisis did not discriminate across gender lines as female wages fell by about the same percentage as male wages. As was true for men, female wages declined more sharply in the urban areas—about 8% more than in rural areas. During the growth years between 1986 and 1997, female wage growth had been 10 to 15% higher in urban wage jobs than in the rural wage jobs.

In contrast with wages, the overall fraction of the population in the labor force changed very little during the decade prior to the crisis and during the crisis. What action there is lies in movement between sectors of employment. During the growth years, male and female employment rates were increasing in the wage sector and declining in unpaid family jobs. The 1986–1997 growth in wage employment was about 3 percentage points higher for men than it was for women while the decline in unpaid family jobs was about two percentage points greater for women compared with men (of whom very few were unpaid family workers). The principal gender difference was that men were moving out of self-employed jobs while this sector remained relatively stable among women. These broad trends suggest that the main employment impact of the crisis was to shift male employment out of the wage sector into self-employment which, we will see below, is a shift that had already begun before the crisis.

For this paper, the key point emerging from Table 1 is that the employment trends and changes in wage rates are virtually identical across all three geographic samples. Put another way, the crisis has been far-reaching and is not concentrated in a small number of provinces—the IFLS2+ sample provides a reasonably accurate portrayal of the labor market consequences of the Indonesian economic crisis. Table 1 indicates that we can safely restrict our analysis of SAKERNAS to those provinces that are also in IFLS2+ as we explore more deeply the nature and extent of the crisis. Thus, all analyses reported below use data only from the seven provinces in IFLS2+.

4 Indonesian Labor Market Before and During the Economic Crisis

What happened in urban and rural Indonesian labor markets as a result of the economic crisis? To answer that question, it is necessary to first establish the salient labor market trends in the years preceding the economic shock and then to isolate those trends that were disrupted. The best data to accomplish this goal are the annual waves of SAKERNAS spanning 1986 through 1998.

Employment Rates

Table 2 presents aggregate labor market data for males and females, separating urban and rural residents. In addition to total labor force participation, the table reports employment rates in three sectors: the formal or market wage sector, self-employment, and unpaid family work. The fifth column in each panel lists mean hours worked per week, conditional on working, and the final column records mean hourly wages for those working in the market wage sector. Figure 2 graphically displays the data on market sector employment (in Panel A) and wages (in Panel B).

While rhetoric on the Indonesian labor market has alleged that the crisis has resulted in massive decreases in employment, or increases in unemployment, the data suggest otherwise. Between the mid-1980s and mid-1990s, male employment rates grew modestly in the urban sector and were relatively stable in the rural sector; there was a slight reduction in employment of males between 1997 and 1998—about two percentage points in urban areas and half that much in rural places. Female employment rates grew by around 25% in the urban sector during the decade prior to the crisis. That trend was reversed with the crisis: between 1997 and 1998, female employment rates declined slightly. In the rural sector, the 1980s and 1990s saw a secular decline in employment. That trend was also reversed with the crisis: the female employment rate expanded by almost three percentage points between 1997 and 1998. This suggests that one way rural families may have coped with this economic crisis is by additional work by women.

The relatively small change in total employment rates of men observed during the crisis does not imply that the disruptions in Indonesian labor markets were small. They certainly were not. Across sectors, male employment dropped by about 3.5 percentage points in the wage sector while increasing in self-employed and family jobs. This loss of male jobs in the wage sector was true in both urban and rural areas. However, it would be an over-statement to attribute the decline entirely to the crisis since market sector work peaked for rural and urban men in August 1995, well before the crisis began. The critical shift in the labor market that occurred contemporaneously with the crisis lies in the industrial distribution of jobs. In the three years prior to the crisis, there had been significant expansions in employment in manufacturing, construction, retail, and sales; all of

Table 2 Employment rates and hours of work, by gender, sector of residence, and year

Sector of residence/Year	*MALE*						*FEMALE*					
	Sector of employment				*Hours*	*Wages*	*Sector of employment*				*Hours*	*Wages*
	Any work	*Wage sector*	*Self-emp*	*Family work*	*per week*	*per hour*	*Any work*	*Wage sector*	*Self-emp*	*Family work*	*per week*	*per hour*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
URBAN												
1986	66.3	39.1	24.8	2.5	46.6	1109	30.5	14.2	10.9	5.4	43.3	684
1987	67.3	39.4	24.8	3.1	46.7	1137	32.4	15.3	11.2	5.9	44	743
1988	65.6	37.8	24.7	3	46.6	1013	33.4	14.9	11.8	6.7	43.5	645
1989	65.2	37.9	24.3	3	47.7	1089	31.7	15.3	10.7	5.7	44.1	719
1990	69	41	24.8	3.2	47.6	1201	35.1	17.7	11.2	6.2	44.4	779
1991	70.7	41.6	26	3.1	48	1189	35	18.7	10.4	6	45.7	764
1992	69.9	40.6	26.6	2.7	47.6	1233	35.6	18.4	11.1	6.1	44.5	845
1993	70.3	41.2	26.5	2.6	47.4	1403	35.1	18.9	10.5	5.7	44.2	962
1994	71.2	42.4	26.3	2.6	46.8	1185	36.5	19.3	11.4	5.9	43.8	829
1995	72.1	43.9	26.3	2	47.1	1406	34.3	19.6	10.9	3.8	43.7	1053
1996	71.8	43.2	26.8	1.8	46.5	1522	36.7	19.4	13	4.3	42.5	1116
1997	71.9	42.6	27.2	2.1	46.8	1578	37.7	19.8	12.5	5.4	42.9	1260
1998	69.7	38.8	28.3	2.5	45.8	1104	37	18.8	12.2	5.9	41.9	835
RURAL												
1986	85.8	22.7	52	11.1	41.1	740	56	9.9	17.2	29	31.7	446
1987	85.9	24.1	50	11.8	41	691	57.2	10.6	16.6	30	31.5	434
1988	85.4	23.9	49.2	12.3	41.4	671	58.4	11.1	16.6	30.7	32.2	394

1989	84.3	23.4	48.7	12.2	41.8	730	58.6	11.4	17.2	30	33.1	429
1990	86.5	23.4	51.2	11.8	41.5	727	55.5	10.5	16.4	28.6	32.1	462
1991	86.2	23.7	51.7	10.8	41.5	759	54.5	10.1	16.2	28.1	32.1	533
1992	86.4	23.4	52	11	40.4	786	56.6	11.3	16.1	29.2	31.4	545
1993	85.7	25.1	51.2	9.5	40.8	866	54.4	11.3	16.7	26.3	32.2	592
1994	85.3	28.1	48.9	8.3	40.8	842	53.6	11.1	18.9	23.6	33	568
1995	84.7	29.3	47.5	7.9	41.4	935	45.2	10.2	17.2	17.8	32.5	710
1996	85.4	27.2	52.2	5.9	39.5	1007	53	11	23.8	18.1	30.7	705
1997	85	27.9	49.8	7.3	40.7	1048	51.7	11.2	18.6	21.9	32	772
1998	84	24.4	51.5	8.1	38.6	745	54.6	11.3	19.9	23.4	30.5	552

Source: SAKERNAS using IFLS2+ provinces only. All hourly wages in 1997 Rp deflated with BPS province-specific price indices. All rates and means weighted

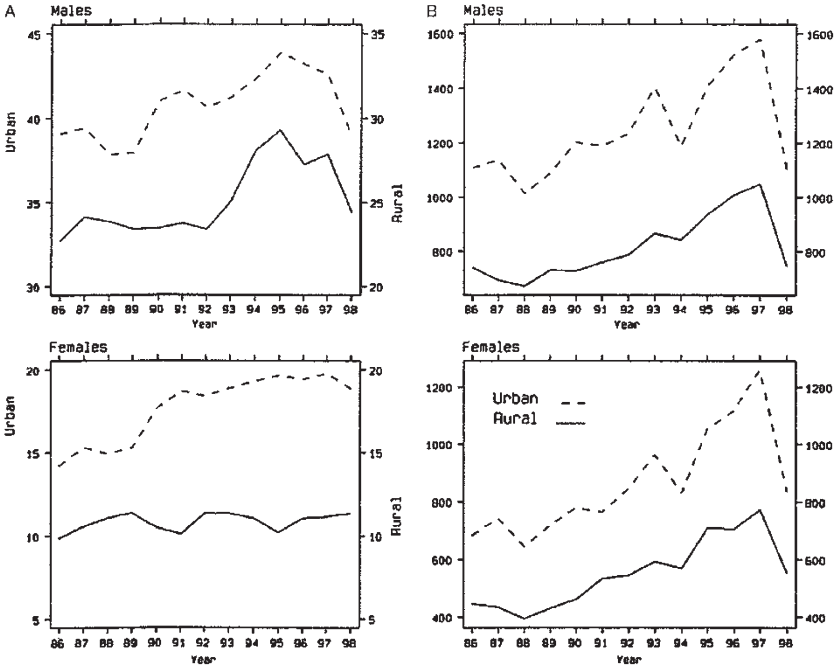

Fig. 2 Employment and hourly wages by year

these sectors saw a dramatic reversal of that trend.[7] Government employment remained relatively stable through the crisis. The absorbing sector was agriculture where employment expanded by 7.2 percentage points (or 20%) between August 1997 and August 1998—a reversal of a secular decline in the sector over the previous decade. This expansion in agricultural work took place in both rural and urban areas.

Employment in the wage sector did not contract for women living in rural areas and did so only slightly for female urban residents. The additional entry of women into the rural labor force is about equally distributed between self-employed and unpaid family jobs. Like men, employment of women grew in the agriculture sector—by about the same amount as

[7] Retail, construction, and manufacturing grew by 25%, 22%, and 11%, respectively, between 1995 and 1997; in construction and manufacturing, all of the gains were lost by 1998; half the gain was lost in the retail sector.

men—and the principal contracting sectors for women were textiles, retail, sales, and services.

Column 5 of each panel reports average hours of work among those working. It has remained reasonably stable from 1986 through 1997 for all workers. Between 1997 and 1998, however, there was a substantial cut in hours worked—about an hour per week in the urban sector and nearly two hours per week in the rural sector.

Wages

Overall, the crisis has had a small impact on aggregate employment. Wages present a sharp contrast: the drama of the crisis is reflected in the collapse of real wages of both men and women. This is vividly displayed in panel B of Fig. 2 which places the real wage cuts between 1997 and 1998 in the context of a substantial rise in real wages during the previous decade. Figure 2 highlights the fact that this wage growth was especially fast in the three years preceding the crisis, particularly in urban areas. Since hours of work fell with the crisis, on average, individual earnings also collapsed during the crisis.

Trends in average wages can mask important changes in the structure of the wage distribution. It has been argued that the crisis in Indonesia has mostly affected the urban elites (Poppele et al., 1999) which would be consistent with a large decline in the mean wage but relatively modest changes in real wages of the poorest. Figure 3 summarizes the entire distribution of wages and how it has changed over time. Non-parametric estimates of the real wage distribution in the formal wage sector for 1986, 1997, and 1998 are presented. The figure demonstrates that during 1998 the *entire* distribution of male and female wages shifted sharply to the left, relative to 1997. For men, this single year shift in the wage distribution virtually wiped out all wage gains made since 1986. Among women, the leftward shift in the distribution is large but they have maintained some of their wage gains since 1986. The key point is that the crisis has affected the entire wage distribution: between 1997 and 1998, the lowest wage earners and the highest wage earners have all seen real wage cuts, on the order of around 40%.

Sub-group Differences

To isolate those sub-groups that have been affected most by the crisis and to assess the statistical significance of the labor market changes, we have

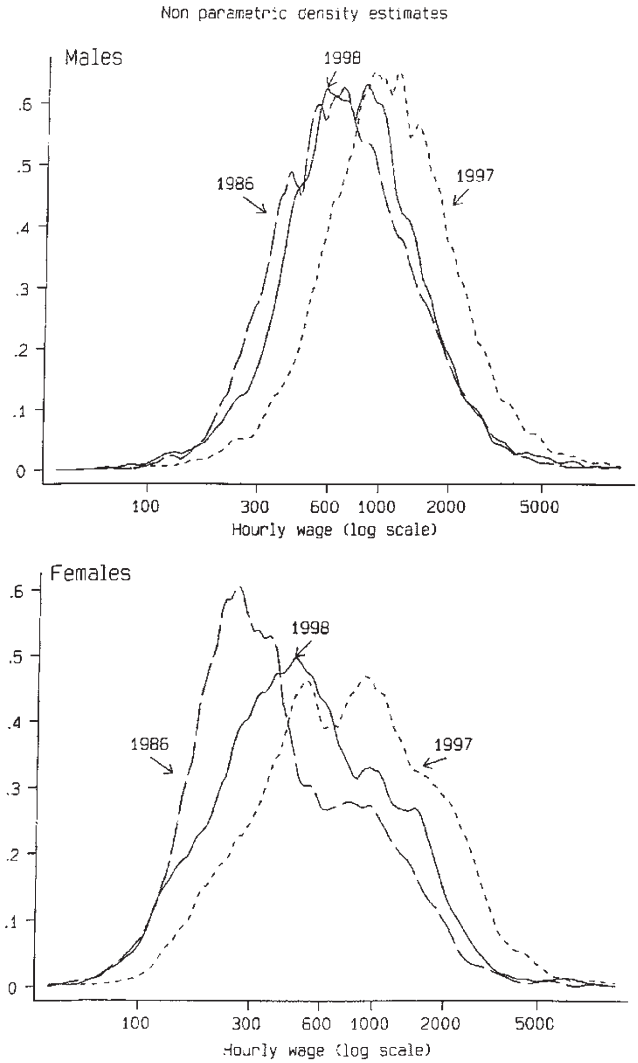

Fig. 3 Non-parametric density estimates of hourly wage distribution: 1986, 1997, 1998

estimated a series of OLS regressions of labor market outcomes. Since employment and wage trends may be affected by shifting characteristics of the labor force, the regressions control province of residence, age (specified as splines with knots at 25, 40, and 65 years), education (stratified into five schooling categories), and urban-rural residence. Thus, the estimated year effects are standardized for changes in the attributes of workers which is potentially important given the changes in employment rates of women over the last 15 years and the substantial upgrading of the education of the labor force during this time. The models are estimated separately for men (in the left panel) and women (in the right panel). Each row of the panel represents a different sample. The first row includes all respondents; we then stratify by sector of residence, by education of the worker, and finally by age. In each stratified model, variables differentiating members of the group are obviously excluded.

Our focus is on changes between 1997 and 1998. The regressions include a calendar year dummy for every year of the survey with 1997 being omitted. Table 3 reports the coefficient on the 1998 calendar year control which measures the average (worker attribute standardized) change between 1998 and 1997 and thus highlights the impact of the economic crisis. Variance-covariance matrices are estimated by the method of the jackknife. Since SAKERNAS contains a large number of observations in each year, conventional rules of thumb for judging statistical significance are likely to be too lax. We adopt a Bayesian approach and, following Schwarz (1978), choose a critical value so that the a posteriori most likely model is selected; in the linear regression case, this amounts to choosing a critical value of the test statistic that is equal to the square root of the logarithm of the number of degrees of freedom multiplied by the number of restrictions. For a t-test on a particular coefficient, in the first row of the table, the critical value is 3.66 (for employment rates) and 3.5 (for wages). We adopt this more conservative approach for judging significance in favor of a classical approach.

Several conclusions about employment can be drawn from the table. First, changes in total employment rates were relatively modest. Among men, controlling observed characteristics, there was a 1.5 percentage point decline in total labor force participation—a decline that is concentrated in urban areas, among more educated workers and among younger workers. There is no significant impact on employment rates overall for women although, after the crisis, rural women, those with elementary

Table 3 Magnitudes of employment and wage shocks

	MALE					*FEMALE*				
	Sector of employment				*Wages*	*Sector of employment*				*Wages*
	Any	*Wage*	*Self*	*Family*	*per hr*	*Any*	*Wage*	*Self*	*Family*	*per hr*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
All country	-1.47	-3.68	1.74	0.47	-38.46	0.81	-0.59	0.5	0.89	-38.99
	(5.7)	(10.87)	(5.31)	(2.81)	(53.17)	(2.24)	(2.26)	(1.84)	(3.4)	(33.42)
Sector										
Urban	-2.74	-4.21	1.08	0.39	-39.99	-1.16	-1.32	-0.35	0.51	-43.28
	(6.94)	(8.39)	(2.39)	(2.25)	(43.71)	(2.26)	(3.22)	(0.99)	(1.99)	(30.28)
Rural	-0.34	-3.27	2.41	0.52	-36.08	2.51	0.07	1.23	1.2	-31.20
	(1.04)	(7.24)	(5.15)	(1.88)	(30.98)	(5.01)	(0.22)	(2.98)	(2.7)	(15.96)
Education										
None	-0.92	-3.05	1.61	0.52	-38.23	1.4	1.21	0.2	-0.02	-30.38
	(0.72)	(2.38)	(0.97)	(0.91)	(7.19)	(1.26)	(1.74)	(0.2)	(0.02)	(6.17)
Elementary	1.32	-1.64	2.77	0.19	-33.06	4.36	0.96	1.42	1.98	-25.12
Incomplete	(2.31)	(1.97)	(3.05)	(0.51)	(14.63)	(5.18)	(1.77)	(1.91)	(2.81)	(7.37)
Elementary	-0.68	-4.02	2.73	0.61	-37.56	2.55	0.57	1.17	0.81	-40.51
Complete	(1.79)	(7.06)	(4.68)	(1.94)	(28.94)	(4.15)	(1.39)	(2.45)	(1.69)	(18.09)
Jr Secondary	-2.26	-3.20	0.5	0.43	-39.87	-1.19	-1.51	-0.03	0.36	-55.37
	(3.76)	(4.42)	(0.73)	(1.02)	(23.17)	(1.47)	(2.72)	(0.06)	(0.65)	(16.64)
Sr Secondary/	-3.40	-4.60	0.72	0.49	-39.44	-4.15	-4.09	-1.24	1.18	-39.46
Tertiary	(6.13)	(6.55)	(1.22)	(1.76)	(33.18)	(4.83)	(4.97)	(2.64)	(2.88)	(23.4)
Age										
15–24	-2.08	-4.02	0.94	1	-38.50	-0.86	-2.02	0.37	0.79	-44.16
	(3.05)	(6.8)	(2.04)	(2.12)	(23.31)	(1.33)	(3.93)	(1.2)	(1.83)	(22.14)
25–39	-1.89	-4.67	2.26	0.52	-41.07	0.66	-0.11	0.02	0.75	-39.68
	(5.57)	(7.43)	(3.68)	(2.26)	(40.4)	(1.07)	(0.24)	(0.05)	(1.67)	(22.69)
40–64	-1.24	-2.74	1.61	-0.11	-33.75	2.27	-0.10	1.32	1.05	-30.71
	(3.71)	(4.19)	(2.41)	(0.79)	(25.4)	(3.21)	(0.22)	(2)	(1.83)	(12.12)
65–75	1.74	-1.73	3.75	-0.28	-31.57	3.66	1.45	0.97	1.24	-13.29
	(1.5)	(2.15)	(3.05)	(0.79)	(5.89)	(2.98)	(2.63)	(0.9)	(1.55)	(1.55)

Regression estimates of differences in year effects between 1997 and 1998 by gender, sector of residence, level of education, and age

Source: SAKERNAS. Based on OLS regressions including individual respondents in all years 1986 through 1998 resident in IFLS2+ provinces. Regressions control province of residence, sector of residence, age and education of respondent. Percentage point change between 1997 and 1998 reported for employment rates; percentage changes reported for wages. Associated *t* statistic reported in parentheses below each estimated change. Variance-covariance matrix estimated by method of jackknife. Estimates for all country (in first row) based on 638,836 observations (male employment), 194,379 observations (male wages), 658,259 observations (female employment), 85,573 observations (female wages). Critical value of t statistic for significance based on Schwarz criterion is 3.66 for employment rates and 3.5 for wage functions in these cases

schooling and older women were more likely to be working while urban, better educated women were less likely to be working.

There was more action in the sectoral distribution of work. Male employment declined by about 3.7 percentage points in the wage sector. This employment reduction was partially absorbed by the total labor force withdrawal mentioned above and partly by an increase in self-employed work and, to a lesser extent, work on the family farm. The reduction in the probability of working in the wage sector was particularly pronounced among younger men; older male workers seemed to have been better able to find self-employed jobs which presumably served to cushion the impact of the crisis.

The increases in employment rates among women discussed above are reflected in increases in self-employment and in unpaid, family work. This is true for rural women, women with elementary schooling, and older women. Running counter to overall trends, employment rates of better educated women declined. For example, female employment in the highest schooling group fell by 4.2 percentage points—an employment decline that is concentrated in the market wage sector.

For both men and women, wages fell by almost 40% between 1997 and 1998. Among male workers, no one has been spared: the wage declines are slightly higher in urban areas, uniformly distributed across education groups and are largest for the youngest. There are larger, and significant, differences in the impact on sub-groups of women. The crisis has affected the wages of urban wage workers more, women in the middle of the education distribution have taken the biggest hit as have younger women. In fact, the only group that has not taken a very large real wage cut are women with wage-earning jobs who are age 65 to 75—a very select (and small) group. In sum, given the magnitude of these wage changes and the relatively small employment effects, the evidence suggests that short-run labor supply functions are fairly inelastic in Indonesia.

The statistical models reported in Table 3 provide estimates of only average effects. How were these real wage declines distributed across lower, middle, and higher wage workers? Table 4 lists the percent real wage declines that took place in the year of the crisis at selected percentiles of the wage distribution in the wage sector. To place these wage decreases in the context of the substantial rise in real wages accompanying economic growth of the previous decade, we also list real wage growth between

Table 4 Distribution of wage shocks

	Urban and rural		*Urban*		*Rural*	
	% change between		*% change between*		*% change between*	
	86 and 97	*97 and 98*	*86 and 97*	*97 and 98*	*86 and 97*	*97 and 98*
	(1)	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*
Wage percentile						
MALES						
10	47.8	-35.9	47.0	-41.6	41.4	-31.6
30	49.0	-35.9	37.9	-39.9	43.0	-33.0
50	43.6	-36.8	31.1	-37.3	43.0	-34.3
70	36.9	-36.7	31.0	-36.4	35.6	-32.9
90	33.5	-36.4	34.4	-36.3	27.1	-40.1
95	37.7	-37.4	38.7	-34.0	20.9	-34.0
FEMALES						
10	50.9	-36.0	68.5	-47.8	34.4	-28.1
30	69.3	-40.6	81.0	-47.1	49.9	-30.3
50	81.0	-46.5	71.7	-50.6	55.1	-31.3
70	77.0	-42.2	50.7	-42.3	58.7	-34.4
90	61.6	-38.3	51.2	-38.1	59.7	-37.7
95	53.6	-38.2	49.0	-37.2	47.4	-35.1

Short-term and longer term percent wage changes across percentiles of wage distribution

Source: SAKERNAS for IFLS2+ provinces. Estimates in column (1) are ln(wage) in 1997 evaluated at appropriate percentile of 1997 wage distribution—ln(wage) in 1986 evaluated at same percentile of 1986 wage distribution. Estimate in column (2) is ln(wage) in 1998—ln(wage) in 1997 evaluated at appropriate percentiles

1986 and 1997 for the same percentiles. The distributions are arrayed separately by gender and by sector of residence.

In many ways, the Indonesian economic crisis was an equal opportunity destroyer. Among men, the substantial fall in real wages during 1998 was only slightly greater among higher wage workers than lower wage workers. This uniformity contrasts with the more beneficial effects of economic growth on male wages at the bottom of the wage distribution between 1986 and 1997. As a consequence, the economic crisis essentially wiped out the previous decade's real wage increases among higher wage men, but at least left lower wage market sector workers somewhat better off than they were in 1986. This story of uniform male wage declines during the crisis has to be amended when the data are stratified by location of employment. Very low wage male workers in urban areas had about a 7%

greater wage fall than high wage urban male workers did. This contrasts with an opposite distributional pattern in rural settings where there was a tendency for higher wage rural workers to experience slightly higher wage declines. In urban and especially rural areas, the prior distributional pattern of economic wage growth differentially favored lower wage workers.

Among female wage earners, the situation was somewhat different as the hardest hit during the crisis were neither very low nor very high wage earners. Rather, those women who suffered the most were in the middle of the wage distribution. They experienced wage declines about 10% larger than those at the bottom or top of the wage distribution. As the accompanying data for 1986–1997 indicate, wages in the center of the distribution grew the most. The concentration of female wage losses in the middle of the wage distribution is a consequence of quite different patterns within the urban and rural sectors. All urban female workers at or below the median experienced about a 10% greater wage loss than urban female workers above the median. Low-wage urban female workers are around median wage earners in the aggregate distribution and so it is this group that is producing the bulge in the wage shock at the center of the aggregate female wage distribution. In contrast with low-wage workers taking the bigger hit in the urban sector, among rural woman, wage earners in the upper part of the distribution experienced bigger wage declines.

A key goal of this research centers around the differential impact of the Indonesian economic crisis on workers in urban and rural labor markets. To gain some insight into this issue, it is not sufficient to simply estimate mean wage effects within each sector. Mean wages and skills are lower in rural areas so that an impact of workers of the same skill would be detected at a higher point in the rural wage distribution than in the urban wage distribution. To see whether economic shocks affect workers of similar skills differentially in rural and urban labor markets, we first arrayed (within gender) the data by percentiles in the aggregate 1997 wage distribution. We next found the corresponding percentiles in the urban and rural wage distribution that matched each aggregate percentile wage. Given that the distribution of urban wages lies above that of rural wages, these matched urban wages were found at a lower percentile and the matched rural wages at a higher percentile than the aggregate percentile wage. For each sector of residence, we then computed the percent change in the wage between 1997 and 1998 yielding a difference in rural wages and a difference in urban wages at the same real wage. The difference between the rural wage change and urban wage change is indicative of what

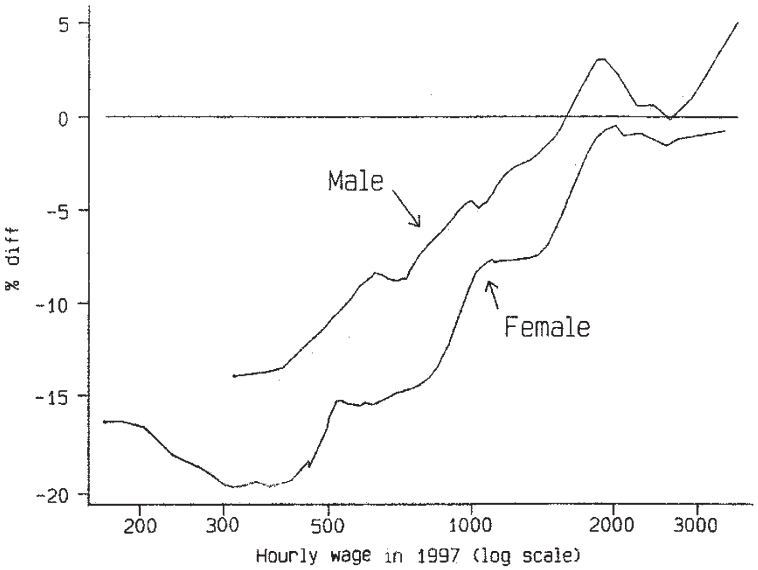

Fig. 4 Urban-rural differential in % wage decline between 1997 and 1998

happened during the crisis to the urban-rural wage differential among workers with the same wage in 1997. If wages are equated with skills, this "difference-in-difference" provides an indicator of the differential impact of the economic shock between urban and rural labor market at each point in the skill distribution.

Figure 4 plots the (smoothed) estimates of the urban-rural differential in the wage decline during the year of the shock separately for male and female wage earners across the 1997 wage distribution. The patterns are remarkably systematic. Among the least skilled men, there was about a 15% decrease in urban wages relative to the drop in rural wages. This greater relative wage deterioration in urban markets monotonically declines as we move up the wage (skill) distribution until there is an equal reduction of male urban and rural wages at the highest wage (skill) levels. The evidence for women essentially parallels that for men: among the lowest wage (skilled) workers, urban wages declined by between 15 and 20% more than rural wages but at the top of the wage (skill) distribution, the rural-urban wage differential did not change.

Based on this evidence, we conclude that at least in the formal wage sector, the Indonesia economic crisis hit unskilled labor markets harder in urban areas than in rural settings. In contrast, there was little geographic differentiation among the highly skilled. This difference between the change in the rural-urban wage gap between the lower skilled and higher skilled provides a "difference-in-difference-in-difference" type of test for the completeness of markets. If markets are complete (or efficient), the difference-in-difference-in-difference should be zero.[8] It is not. This is *prima facie* evidence in support of the view that markets are relatively efficient among the higher skilled but that, at least in times of substantial economic turmoil, markets seem to behave as if they are incomplete among the least skilled. It may reflect barriers to mobility from the urban to rural sector among the least skilled (possibly because of liquidity constraints) or incomplete information on the part of the least skilled. In either case, the interpretation is consistent with evidence on the completeness of markets in equilibrium (Pitt & Rosenzweig, 1986; Benjamin, 1994) but incompleteness in the presence of unanticipated shocks (Murrugarra, 1998).

5 The Dynamics of the Labor Market Adjustment

The previous section provided a portrait of the Indonesian labor market before and in the midst of the economic crisis based on SAKERNAS, a time series of cross-sections. While informative, these portraits are inherently limited in two key dimensions. First, SAKERNAS records income only for workers in the formal wage sector; this limitation excludes more than half of the labor force. Second, on the employment side, SAKERNAS is not informative about the dynamics of any labor market adjustments that took place as workers switched their sectoral or geographic places of work. To address these questions, we now turn to linked micro data from IFLS2 and IFLS2+. This panel survey contains information about individuals who were interviewed immediately before the onset of the 1998 Indonesian economic crisis and again a year later.

[8] Prices are probably lower in rural areas. However, uniformly lower prices in rural areas would only alter the location and not the shape of the curves presented in Fig. 4 and so would not affect inferences based on the difference-in-difference-in-difference.

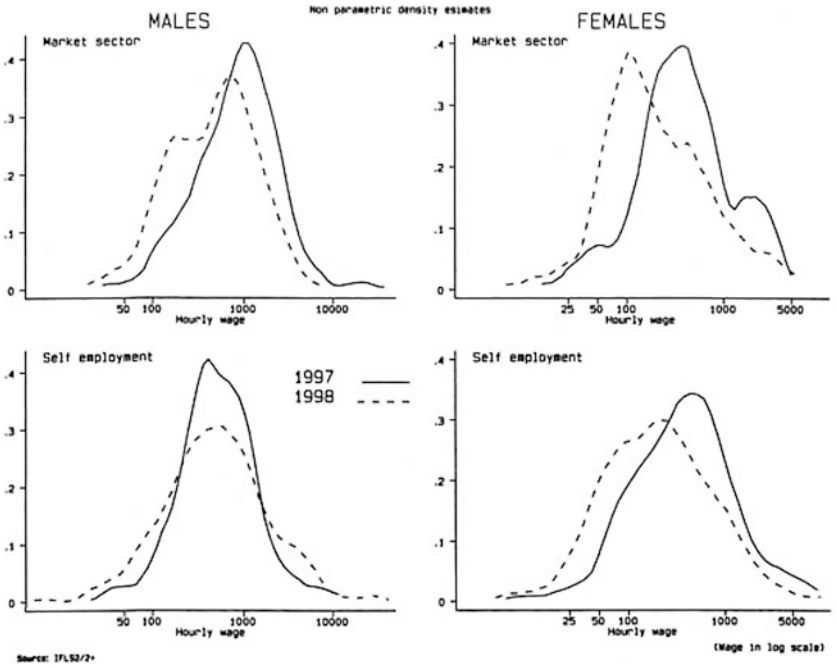

Fig. 5 Non parametric density estimates of hourly earnings distribution: market and self-employed sectors

Market Sector and Self-employed Hourly Earnings

It is well-known that collection of income from self-employment in a survey setting is extremely difficult—especially, perhaps, in a low-income and substantially agricultural setting like Indonesia. Difficulties arise, first, because of the need to calculate costs and net those out to compute profit, and, second, because incomes tend to be volatile over time, often containing an important seasonal component. The panel feature of IFLS provides some assistance. To the extent that the difficulties in measurement for a particular individual do not change between 1997 and 1998, these concerns will be somewhat mitigated; inferences based on changes in self-employment incomes over the period may not be as seriously contaminated as inferences about levels in incomes.

Figure 5 exploits information on hourly earnings in the market wage and self-employed sector reported in both IFLS2 and IFLS2+.

Non-parametric estimates of the distribution of hourly earnings are reported for men and women in each sector of employment separately for 1997 and 1998.

Among men and women working in the market wage sector, the IFLS data confirm the general finding in SAKERNAS that there was a dramatic shift to the left in the distribution of real wages between 1997 and 1998. IFLS indicates there was a similar deterioration in women's income in the self-employed sector. However, in sharp contrast, real median hourly earnings of self-employed men did not change at all during the economic crisis. Clearly, inferences about the impact of the crisis based exclusively on market sector wages (as reported in SAKERNAS) miss part of the picture.

The robustness of self-employment hourly earnings may reflect an increase in the amount of capital used in self-employment businesses (and failure of the respondents to incorporate the costs of capital in their estimates of profits). This seems unlikely given that credit markets were extremely tight at this time. It may also reflect increased labor supply of other family members when self-employment income is derived from family businesses. This is because income from the business is attributed to one family member and the estimated "hourly" earnings do not take into account the work effort of others.[9]

The distribution of hourly earnings in self-employment is fatter tailed than the distribution of wages. In part, this reflects greater error in measurement of self-employed earnings. However, Fig. 5 indicates that, relative to 1997, the distribution of hourly earnings in 1998 among self-employed men was fatter tailed. This increase in variance between the years suggests that some self-employed workers benefitted from the economic crisis in terms of hourly earnings—presumably a reflection of increases in the prices of products they sold—while others appear to have been all but completely wiped out. In the rural sector, more than half of self-employed men actually had positive real income changes during the crisis. Since self-employed men constitute more than half the male work

[9] This assumes only one family member is productive in the business. We have confirmed that the results presented below are not driven by that assumption by taking the opposite extreme and treating all family workers as equally productive. In that case, hourly earnings in a family business are computed as the ratio of total earnings to total number of hours worked by all family members and the same hourly earnings rate is attributed to each family worker. See Thomas et al. (2000a).

force, any characterizations of the economic shock in Indonesia using only the wage sector are both incomplete and misleading.

Before comparing wage and employment changes between the SAKERNAS and IFLS, a few additional adjustments are necessary to make them directly comparable. These adjustments are required because the IFLS is a panel survey while SAKERNAS is cross-sectional. This distinction mainly affects labor force patterns for either very young individuals (who are likely to enter the labor market as they age) or very old individuals (who are likely to exit as they age). To avoid contaminating the comparisons with these effects of aging, we limit our comparisons in this section to individuals who were between the ages of 22 and 64 in 1997. Since IFLS2 and IFLS2+ weights are not yet available, a more reliable basis for a comparison across the two surveys can be obtained from regression models which include as co-variates controls for the main attributes that determine the sampling weights—province, rural-urban residence, and age of respondent. The regression model adjusted impacts reported in

Table 5 Adjusted hourly earnings shocks in market and self-employed sectors

Quantile	*Male*				*Female*			
	Urban		*Rural*		*Urban*		*Rural*	
	Market	*Self*	*Market*	*Self*	*Market*	*Self*	*Market*	*Self*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Mean	-53.5	-66.9	-57.1	-11.2	-58.6	-51.2	-58.6	-55.8
	[9.5]	[5.3]	[6.8]	[1.2]	[6.2]	[3.8]	[4.5]	[4.8]
25%ile	-57.5	-89.0	-59.7	-24.9	-66.5	-76.7	-76.1	-63.3
	[6.4]	[5.2]	[5.6]	[1.8]	[4.4]	[3.6]	[4.5]	[3.3]
50%ile	-46.0	-57.7	-49.8	-10.6	-55.9	-56.0	-65.7	-48.8
	[6.7]	[3.4]	[3.6]	[1.1]	[5.2]	[3.7]	[4.4]	[3.1]
75%ile	-40.3	-45.2	-46.6	1.2	-54.0	-60.8	-48.1	-53.6
	[6.4]	[2.5]	[5.2]	[0.1]	[5.8]	[3.6]	[2.9]	[3.6]
# obs	773	390	486	611	326	358	210	458

Estimated percentage change in earnings per hour between 1997 and 1998, IFLS panel respondents

Source: IFLS2 and IFLS2+ panel respondents. Quantile regression and OLS regression estimates based on panel respondents interviewed in IFLS2 (1997) and IFLS2+ (1998). Estimated percentage difference in hourly earnings in 1998 and hourly earnings in 1997 reported in table; [t statistics] in parentheses estimated by method of jackknife (for mean) and bootstrap (for quantiles). Regressions control age, education, and province of residence of respondent

Table 6 Employment shocks

	All	*Market*	*Self-*	*Family*
	workers	*sector*	*employed*	*workers*
	(1)	*(2)*	*(3)*	*(4)*
	MALES			
URBAN				
SAKERNAS	-2.9	-4.3	1.2	0.2
	[6.4]	[6.9]	[2.1]	[1.3]
IFLS	-1.2	-3.9	1.5	1.2
	[0.7]	[1.6]	[0.6]	[2.0]
Difference	1.7	0.4	0.3	1.0
(IFLS-SAKERNAS)	[1.0]	[0.2]	[0.1]	[1.6]
RURAL				
SAKERNAS	-0.6	-3.3	2.3	0.4
	[2.1]	[6.1]	[4.1]	[1.7]
IFLS	-3.4	-0.6	-1.6	-1.3
	[3.0]	[0.3]	[0.7]	[1.4]
Difference	-2.8	2.8	-3.9	1.7
	[2.4]	[1.3]	[1.7]	[1.8]
	FEMALES			
URBAN				
SAKERNAS	-0.5	-0.5	-0.5	0.5
	[0.8]	[1.0]	[1.0]	[1.5]
IFLS	6.6	1.0	1.5	4.2
	[2.8]	[0.5]	[0.7]	[3.9]
Difference	7.1	1.4	1.9	3.7
	[2.9]	[0.7]	[0.9]	[3.3]
RURAL				
SAKERNAS	3.1	0.3	1.5	1.3
	[5.4]	[0.8]	[3.0]	[2.4]
IFLS	7.0	1.1	0.3	5.7
	[3.4]	[0.8]	[0.1]	[3.1]
Difference	3.9	0.8	-1.2	4.4
	[1.8]	[0.5]	[0.7]	[2.3]

Adjusted differences between 1997 and 1998, SAKERNAS and IFLS

Source: SAKERNAS and IFLS. OLS regression estimates of difference between 1997 and 1998; asymptotic t statistics in parentheses estimated by method of jackknife. Regressions include controls for age, education, and province of residence

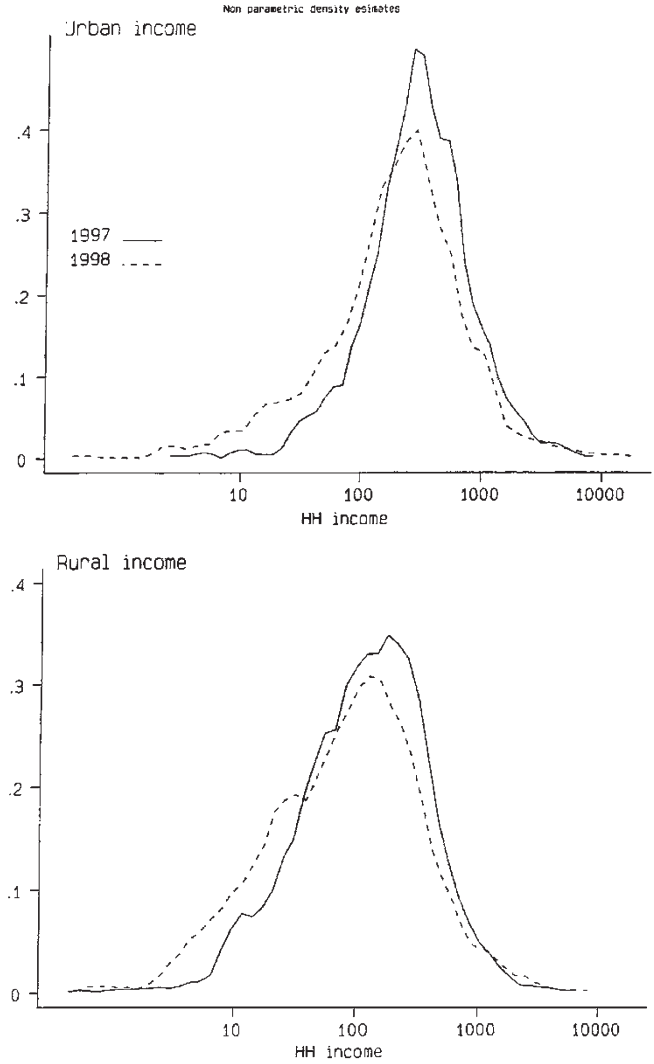

Fig. 6 Non-parametric density estimates of distribution of family income, 1997 and 1998

Tables 5 and 6 will be less influenced by different sample weights in the two surveys.

Table 5 examines changes in wages in more detail. Drawing on all respondents in IFLS who were working in either 1997 or 1998, the logarithm of wages has been regressed on an indicator variable for 1998, controlling province of residence, age, and education (all measured in 1997). The regressions are stratified by gender, sector of residence, and sector of work. For each regression, the table presents the coefficient on the 1998 indicator variable which is an estimate of the percent change in real wages between 1997 and 1998 adjusting for inflation and observed worker characteristics. Mean effects, based on OLS regressions, are reported in the first row. To examine the distribution of changes in wages, quantile regressions estimates for the 25th, 50th, and 75th percentiles are reported in the following rows; t statistics are calculated by bootstrapping the quantile regressions and jackknifing OLS. (Since the jackknife is a linear approximation to the bootstrap, the two are asymptotically equivalent in the linear regression case.)

The IFLS percentage wage declines in the market sector are larger than those summarized in the previous section using SAKERNAS. In part, this reflects differences in the timing of the surveys. SAKERNAS is conducted in August each year; most of the IFLS interviews were conducted between August and December in each year. Real wages were relatively constant in 1997. Evidence from published sources suggests that in 1998 real wages continued to decline from August until December (see, e.g., Papanek & Handoko, 1999). Thus, the IFLS wage changes should be larger than those in SAKERNAS.[10]

Figure 5 indicated that self-employment hourly earnings of males were remarkably robust during the crisis. Table 5 reveals that this is confined to males working in the rural sector. Among urban men and all women in the IFLS, self-employed hourly earnings declined as much as market sector wages.

Among workers in the market wage sector, percentage declines in wages are slightly higher in the bottom quartile of the wage distribution. In contrast, there exists a much larger negative decline at the first quartile of the self-employed wage distribution compared to the third quartile

[10] If interviews were randomly allocated over time in IFLS, we could control month of interview and directly compare IFLS with SAKERNAS. The interviews are not randomly assigned. Because IFLS is a panel survey, interviewers begin in areas that tend to be places migrants move from and end in areas that are destinations for migrants. This reduces the costs of tracking movers. Therefore, we do not attempt to extract any information from the month to month differences within IFLS.

especially among men. It is common to observe that self-employment hourly earnings are more dispersed than market sector wages; not only is this true in the IFLS but, more significantly, the difference in dispersion is greater after the crisis than before.

While the regressions reported in Table 5 estimate the degree of shift at different points in wage distributions between 1997 and 1998, they do not speak directly to the distributional impacts of the crisis since it is not describing wage changes for individual workers. Wages at the first quartile of a wage distribution could decline more than wages at the third quartile for several other reasons. For example, if relatively high-wage rural workers in the wage sector moved to low-wage employment in the rural self-employed sector, the rural sector wage distribution would shift downward less at the top of the rural wage distribution. A more direct test of the distributional impact of the crisis on wage rates involves re-estimating the models described in Table 5 and testing whether percent wage reductions in 1998 varied by education level. For the combined market and self-employed wage distributions, we found that there was no systematic pattern across respondents across the education distribution. This suggests that based on wages alone, the impact of the crisis was fairly uniform across low- and high-skill workers.

Employment Rates and Transitions

We return to the employment side of the labor market. Table 6 presents estimates of changes in sector-specific employment rates between 1997 and 1998 for IFLS and SAKERNAS. The estimates in the first column are based on OLS regressions of the probability a respondent is working in either year controlling province of residence, age, education, and an indicator variable for 1998. The estimate on the 1998 indicator variable is reported in the first and second row of each panel for SAKERNAS and IFLS, respectively. The difference between the IFLS and SAKERNAS estimates is reported in the third row. The remaining three columns are based on regressions of the probability a respondent is working in a particular sector. The estimates based on IFLS are less precise than SAKERNAS because the IFLS samples are considerably smaller.

For men, the results parallel those discussed above. There are small changes in labor force participation rates of urban males with declines in activity in the market sector offset by increases in self-employment and family work. None of the differences between SAKERNAS and IFLS is significant. Among rural males, IFLS apparently overstates the decline in participation rates relative to SAKERNAS.

There is one major difference in the results for women. Relative to SAKERNAS, IFLS estimates significantly larger increases in labor force participation rates during the crisis. The difference is concentrated in a greater increase in unpaid family work in IFLS. In both urban and rural areas, there is over a four percentage point increase in the fraction of women working as unpaid family workers during the economic crisis suggesting that one way households have adapted to the economic shock has been a significant expansion of work activity by women in family businesses.

In economies such as Indonesia where formal wage sectors co-mingle alongside more traditional economic activities that often take place within the family unit, measuring work activity in surveys is not straightforward. SAKERNAS and IFLS take a somewhat different approach to this issue. Both surveys begin with the same lead-in question on whether, during the week prior to the survey, the respondent was working.[11] In SAKERNAS, those who answer "yes" to that question are then asked their type of work. Family work is one option. The same screener question is asked in IFLS and those who answer "yes" are asked about the type of work. The difference between SAKERNAS and IFLS lies in the treatment of those who answered "no" to the lead-in question. In SAKERNAS, they are not asked any more questions about work. In IFLS, these respondents are asked a follow-up question about whether they had worked for one hour for pay and, if the response is no, then they are probed by a second follow-up question about whether the respondent had worked in a family-owned business. Those who answer "yes" to either the lead-in or the two follow-up questions about family business are treated as working and complete the employment module. This additional probing in IFLS should lead to higher estimates of labor force participation rates especially among women (which it does). But more important, this additional IFLS probing should make it easier to detect individuals switching into the family employment sector. If the increase in family work reported in IFLS arises because of the probing, the follow-up question should have captured more positive responses in IFLS2+ relative to IFLS2. It did. In IFLS2, 1.8% of female respondents who answer the question said they worked in a family business; in IFLS2+, 4.1% of women answered affirmatively. This suggests that many of the people who have entered the family business would not ordinarily consider this activity as work in some traditional sense. Given this

[11] Work is defined as an activity for the sake of generating income or helping to generate income that takes up at least one continuous hour.

evidence, we are inclined to find the IFLS results on female family work more credible than those in SAKERNAS.[12]

To this point, the discussion of labor force activity has been based on cross-sectional evidence. Table 7 makes use of the repeated observations on individuals in IFLS and tabulates transition rates between 1997 and 1998 for each type of work activity, conditioning in each case on the respondent's activity in 1997. The relatively stable aggregate data on male work activity derived from the cross-section comparisons mask considerable movement among the sectors and between work and non-work during the crisis. This churning is particularly prevalent in rural areas where one-third of the men and half the women changed their sector of work between 1997 and 1998.

Among rural men, the largest net transition that took place was a flow of workers from the formal market sector to self-employed jobs. In total 29% of rural men in wage jobs in 1997 moved to the self-employed sector by 1998. The comparable reverse flow was considerably smaller with only 15% of those in self-employment in 1997 working in the wage sector in 1998. This net flow of male employment from the wage to the self-employed sector during 1998 is consistent with our earlier finding that, in rural areas, hourly earnings of the self-employed decreased much less than of those in wage sector jobs. Among rural men, there was also a substantial flow from family work into self-employment as more than half of those rural men who were family workers in 1997 were self-employed by 1998. Note as well that there was much less movement for men between the self-employed and wage sectors in urban areas where the wage decreases in each sector were more similar.

Among women, the most salient pattern is the significant inflows into unpaid family jobs from all other sectors. In rural areas, the fraction of women working in family jobs increased from 22 to 27% between 1997 and 1998. Table 7 indicates that these net inflows into family jobs drew women from all sectors—those not in the labor force, in the market, and in the self-employed sectors. Although this net inflow into family work also characterized urban women, rates of employment in family jobs remain much lower there than in rural areas.

[12] See Korns (1987) who argues that there is evidence in SAKERNAS that self-employment and family employment do not seem to mean the same thing to all interviewers. In addition, he argues that part of the year to year variation in self-employment and family employment in SAKERNAS reflects changes in interviewer quality and training over years.

Table 7 Transition matrix

	1998									
	Males					*Females*				
1997	*Market sector*	*Self-employed*	*Family business*	*Not working*	*Σ*	*Market sector*	*Self-employed*	*Family business*	*Not working*	*Σ*
URBAN										
Market sector	81	9	2	8	*57*	76	9	3	13	*22*
Self-employed	14	79	2	5	*30*	6	68	9	17	*23*
Family business	17	50	17	17	*1*	11	7	68	14	*3*
Not working	22	18	2	58	*12*	8	13	5	74	*52*
Σ	*54*	*32*	*2*	*13*	*100*	*22*	*24*	*7*	*46*	*100*
RURAL										
Market sector	63	29	3	5	*32*	53	9	14	24	*11*
Self-employed	15	76	2	7	*58*	6	60	20	14	*27*
Family business	18	51	24	7	*5*	4	17	48	31	*22*
Not working	29	19	12	40	*5*	9	16	25	50	*40*
Σ	*31*	*57*	*4*	*8*	*100*	*12*	*27*	*27*	*34*	*100*

Inter-sectoral employment transitions between 1997 and 1998

Source: IFLS2/2+ panel respondents

We have also examined which type of workers were more mobile between employment sectors. On this issue, we found that by far the least mobile workers were high-wage rural workers in the formal wage sector. This is particularly clear for men.

Why are these workers so reluctant to change the sector of their employment in spite of the large decreases in their wages and the possibility of better opportunities in other sectors? The answer appears to be that they work for the government. In rural areas, two-thirds of men in the upper quartile of the market wage distribution work for the government. The comparable fraction for women is more than three-quarters.

Family Income

To this point, we have analyzed Indonesian labor markets from the perspective of the individual worker. But perhaps even more so than in developed countries, the family may represent a more appropriate unit in which to measure how an economic crisis has impacted the well-being of individuals. The family represents an important institution that may help smooth out economic shocks among its members and thereby mitigate some of the risk associated with the crisis. The results obtained thus far suggest that such adjustments may well be taking place in Indonesia. The increase in female work in family businesses may be seen as a reaction to the poorer labor market prospects of men during the crisis. If so, the impact of the crisis on family earnings may be smaller than that on earnings of individual workers.

To examine this issue, Table 8 presents estimates of changes in family earnings between 1997 and 1998.[13] The left hand panel of Table 8 contains estimates for family earnings as conventionally measured while the right hand panel augments family earnings to include the value of own-production of food.

The data summarized in the table indicate that there were considerably different impacts of the economic crisis on family earnings than that observed for individual hourly earnings (and, therefore, total earnings since, recall, hours of work declined during the crisis). This is especially

[13] Household earnings are the sum of individual earnings among all members of a household in 1997; thus, we ignore household splits, joins, and new entrants. We do this in order to focus on the household unit as it existed in 1997, independent of the current location of its members.

Table 8 Family income shocks

	Family earnings from market and self-employment work		*Family earnings and imputed value of home production*	
	Sector of residence in 1997			
	Urban	*Rural*	*Urban*	*Rural*
Mean	-47	-37	-43	-21
	[10.4]	[8.0]	[8.8]	[5.0]
Quantile				
25%ile	-55	-53	-64	-33
	[7.0]	[8.4]	[8.6]	[6.0]
50%ile	-34	-32	-39	-22
	[10.1]	[4.4]	[9.2]	[4.2]
75%ile	-23	-20	-23	-20
	[5.0]	[4.0]	[5.4]	[5.4]
90%ile	-17	-5	-22	-14
	[2.5]	[0.6]	[2.3]	[2.1]

Estimated percentage change in family income between 1997 and 1998, IFLS panel households

Source: IFLS2/2+ panel households. Estimates of percentage decline in family income between 1997 and 1998 derived from OLS and quantile regressions; [*t* statistics] in parentheses

clear in the rural areas. For example, Table 5 indicates that mean wages in the formal sector fell by about 60% in rural areas. The data in Table 8 indicate that mean conventional family income fell by 37% while mean family income including home production declined by 21%. While this latter decline is definitely not trivial—it is only a third of the estimated fall in market wages in rural areas. Evaluated at the mean, the differences between the impact of the crisis on individual earnings and on family earnings are considerably smaller in urban areas. In those areas, at the mean, market wages fell by about 55% while family income (including the imputed value of home production) fell by 43%.

Table 8 also provides evidence on how the shape of the income distribution has changed. Recall the change in the wage distribution is dominated by a shift in location. In contrast, both the location and scale of the family income distribution have changed as reflected in the greater decline in the lower quartile, relative to the upper quartile of the distribution.

Why are our estimates of the impact of the Indonesian crisis so different for family income than for market wages in the formal wage sector? Our analysis points to two principal factors—the lack of any significant impact

Table 9 Family income shocks and family employment

	# workers in each year:	*Market sector*		*Self-employment*		*Unpaid family bus*	
	Yr = 1998		**1998*		**1998*		**1998*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
URBAN							
Mean	-50.9	44.1	-5.8	24.4	3.3	-6.3	22.5
	[4.0]	[11.0]	[0.9]	[5.2]	[0.4]	[0.6]	[1.2]
25%ile	-73.3	44.6	1.6	22.3	-2.7	4.0	-8.1
	[3.0]	[7.5]	[0.1]	[3.5]	[0.3]	[0.3]	[0.4]
50%ile	-36.7	41.5	0.7	20.2	-6.8	-3.4	2.8
	[2.3]	[9.6]	[0.1]	[3.5]	[0.6]	[0.3]	[0.2]
75%ile	-19.5	44.6	-10.3	24.5	3.7	-11.6	26.1
	[1.3]	[10.9]	[1.6]	[3.6]	[0.3]	[0.7]	[1.0]
90%ile	-12.3	36.5	-15.1	24.7	18.4	-18.2	92.8
	[0.5]	[5.8]	[1.4]	[2.4]	[1.2]	[0.6]	[2.0]
RURAL							
Mean	-53.7	53.4	-14.1	19.4	10.8	-5.6	27.6
	[3.5]	[8.9]	[1.6]	[3.2]	[1.2]	[0.9]	[2.9]
25%ile	-85.3	58.8	-5.1	19.7	10.1	-6.2	33.6
	[3.9]	[5.9]	[0.4]	[2.4]	[0.7]	[0.8]	[2.9]
50%ile	-40.8	54.1	-18.7	19.0	2.0	-1.6	21.3
	[1.7]	[8.8]	[1.9]	[2.8]	[0.2]	[0.2]	[1.7]
75%ile	-33.9	41.6	-10.6	11.7	11.9	-1.9	14.7
	[2.2]	[5.4]	[1.1]	[1.5]	[1.1]	[0.2]	[1.4]
90%ile	-8.2	42.4	-27.4	18.4	9.3	-7.3	14.2
	[0.3]	[3.6]	[1.5]	[1.9]	[0.6]	[1.1]	[0.8]

Correlations between number of family members employed in each sector and family income. Main effects and interactions with control for 1998, IFLS panel households

Source: IFLS2/2+ panel households. Estimates based on regressions of ln(family income) on indicator variable for 1998 (column 1), number of workers in market sector, self-employment, and family work in each year (even columns) plus interaction between the number of workers in each sector and 1998 indicator variable (odd columns). [t statistics] in parentheses based on jackknife for OLS and bootstrap for quantile regressions

of mean incomes of the self-employed in rural areas and the insurance role of the family in mitigating the full fury of this economic shock. With regard to the latter, two mechanisms suggest themselves. First, households may exploit the economies of scale associated with combining (or not splitting up). There is evidence that this is important. Mean household size increased by about 5% (Frankenberg et al., 2000). Second, household members may change their allocation of times. Table 9 investigates the

issues further by summarizing estimates obtained from a descriptive model of the links between (the logarithm of) family income and the number of workers in the family.

We have estimated regressions of family income on the number of family members working in the market, self-employed, and family business sector in that year; each family contributes two observations and both income and the number of workers are time-varying. To highlight differences associated with the crisis, the regressions include a 1998 year effect (reported in the first column of the Table 9). The number of workers in each sector is interacted with the 1998 indicator variable and those interactive effects are reported in the odd columns of the table. Column 2, for example, indicates the magnitude of the correlation between the number of family members in the market sector in 1997 and family income in 1997. Column 3 provides the difference between this correlation and the correlation between the number of market sector workers in 1998 and income in 1998. The sum of columns 2 and 3 is the correlation between number of workers in 1998 and income in 1998. Quantile models are also reported to provide a sense of how these effects differ across the income distribution.

Three main points emerge from the table. First, not surprisingly, before the crisis, the ranking of the most important contributions to family income are market wage earners, self-employed, and finally family workers (who contributed not at all on average). Second, the principal impact of the crisis was to diminish the relative importance of wage earners while simultaneously increasing the importance of the self-employed and family workers. The greater role of family workers is particularly pronounced and, at the mean, it is significant in the rural sector.

Third, these differences vary both across sectors of employment and over the distribution of income. In the urban and rural sector, between 1997 and 1998, the correlation between family income and the number of market workers changed little at median income and below but was cut in half at the top of the income distribution. The reverse is true for self-employed workers in urban areas and the change in correlation is consistently positive across the entire income distribution in rural areas. With regard to the number of family workers, the 1998 mean effects are roughly the same in rural and urban areas (being associated with about a 20% increase in income). However, in urban areas, additional family workers make a large difference at the top of the income distribution compared to their correlation with family income at the bottom. For example, we

estimate a statistically insignificant negative correlation at the 25th percentile compared to a large, positive, and significant correlation at the 90th percentile. In urban areas, therefore, the mitigating impact of family workers is most pronounced towards the top of the income distribution. In rural areas, the correlation between income and the number of family workers was quite different: the mitigating impact of family workers is more pronounced in the bottom half of the income distribution.

It is important to recognize that the quantile regression estimates are informative about the marginal contributions of market, self-employed, and family workers at each point in the family income distribution in 1997 and 1998. They do not speak to the question of whether these workers were more (or less) important contributors to income among lower (or higher) income families. Estimation of the same models, stratifying on level of education of the household head, provides some insights into that question and indicates two aspects in which the effect of the crisis varied between poorer and less poor households. In the urban sector, the contribution of market workers to family income in 1998 was reduced far more in households whose heads had no education. In the rural sector, in 1997, the contribution of family workers was greatest among the least educated. Between 1997 and 1998, their contribution increased most among households whose heads had no schooling and those whose heads had more than primary education. As a result, in 1998, among families with little education, family workers, the self-employed, and market workers were about equal contributors to family income.

6 Conclusion

It has been argued that the large and abrupt economic decline in Indonesia has resulted in unemployment on a massive scale. The evidence simply does not support this contention and estimates of unemployment appear to have been grossly exaggerated. In fact, labor markets in Indonesia appear to be remarkably flexible.

The drama of the financial crisis lies not in aggregate employment but in real wages, particularly in the market wage sector. While there have been increases in nominal market sector wages since the onset of the crisis, these increases have been dwarfed by the rapid increase in prices. According to SAKERNAS, real wages have collapsed by around 40% between August

1997 and August 1998. However, estimation of the exact magnitude of the decline in real wages is fraught with difficulties, some of which revolve around computing the appropriate rate of price increases. Inflation has spiraled during this time, there is evidence of regional heterogeneity in price increases and there have been considerable shifts in relative prices.

Mindful of these concerns, there is evidence in SAKERNAS indicating wage cuts have been greater in the urban sector, on average. It turns out that this difference is entirely concentrated at the bottom of the wage distribution where the rural-urban wage differential has declined by around 15%. At the top of the wage distribution, cuts in real wages are about the same for urban and rural workers. This is *prima facie* evidence for flexible (complete) markets among the better skilled but rigidities (incomplete) markets among those with less skills, at least during times of economic disruption.

These facts emerge in two independent series of household surveys—SAKERNAS and IFLS. SAKERNAS, however, does not record information on income from self-employment. IFLS does. Those data indicate that real hourly earnings of the self-employed in urban areas and of females in rural areas have declined by at least as much as real market sector wages. In stark contrast, real hourly earnings of self-employed males in rural areas have remained essentially stable. Given that over half the male work force are self-employed in Indonesia, conclusions about the effects of the crisis on wages that focus only on the market wage sector will substantially overstate the magnitude of the crisis.

In fact, Indonesian families have displayed a remarkable capacity for resilience in the face of the crisis. We estimate that declines in real family incomes are about half the magnitude of the declines in individual hourly earnings. To be sure, this should not be taken to imply the crisis has not been large and serious. It has. Our results do suggest, however, that households have adopted an array of strategies in an effort to mitigate the effects of the crisis.

References

Benjamin, D. (1994). Household Composition, Labor Markets and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica, 60*(2), 287–322.

Cameron, L. (1999, April). Indonesia: A Quarterly Review. *Bulletin of Indonesian Economics Studies, 35*(1), 3–40.

Corsetti, G., Pesenti, P., & Roubini, N. (1998a). What Caused the Asian Currency and Financial Crisis: Part I, A Macroeconomic Overview. Mimeo.

Corsetti, G., Pesenti, P., & Roubini, N. (1998b). What Caused the Asian Currency and Financial Crisis: Part II, The Policy Debate. Mimeo.

Fallon, P., & Lucas, R. E. B. (1999). Economic and Financial Crises: A Survey of the Evidence. *World Bank Research Observer, 17*(1), 21–45.

Feridhanusetyawan, T. (1999). Social Impacts of Indonesia's Economic Crisis: Employment, Income and Poverty Issues. In *Social Impacts of the Asian Economic Crisis in Thailand, Indonesia, Malaysia and the Philippines.* Thailand Development Research Institute.

Frankenberg, E., Beegle, K., & Thomas, D. (1999). The Real Costs of the Indonesia' Economic Crisis: Preliminary Findings from the Indonesia Family Life Surveys. RAND DRU-2064-NIA/NICHD.

Frankenberg, E., Smith, J. P., & Thomas, D. (2000). Economic Shocks, Wealth and Welfare. Mimeo.

Frankenberg, E., & Thomas, D. (2000). Indonesia Family Life Survey: Study Design and Results from Waves 1 and 2. DRU-2238/1, RAND, Santa Monica.

ILO. (1998). Employment Challenges of the Indonesian Economic Crisis. Mimeo, Jakarta.

IMF. (1999). *World Economic Outlook: International Financial Contagion*. World Economic and Financial Surveys, IMF.

Korns, A. (1987). Distinguishing Signal from Noise in Labor Force data for Indonesia. Development Studies Project Research Paper #1. Jakarta: Bappenas/BPS.

Krugman, P. (1998). What Happened to Asia. Mimeo.

Levinsohn, J., Berry, S., & Friedman, J. (1999). Impacts of the Indonesian Economic Crisis: Price Changes and the Poor. NBER Working Paper 7194.

Murrugarra, E. (1998). Completeness of Markets and Economic Shocks. Mimeo, GRADE, Lima.

Papanek, G., & Handoko, B. S. (1999). The Impact on the Poor of Growth and Crisis: Evidence from Real Wage Data. Conference on 'The Economic Issues Facing the New Government', Jakarta, August.

Pitt, M., & Rosenzweig, M. (1986). In I. Singh, L. Squire, & J. Strauss (Eds.), *Agricultural Household Models: Extensions, Applications and Policy.* Johns Hopkins.

Poppele, J., Sumarto, S., & Pritchett, L. (1999). Social Impacts of the Indonesian Crisis: New Data and Policy Implications. World Bank, SMERU Report, Jakarta.

Radelet, S., & Sachs, J. (1998). The Onset of the East Asian Financial Crisis. Mimeo.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics, 6*, 461–464.

Thomas, D., Beegle, K., & Frankenberg, E. (2000a). Labor Market Transitions of Men and Women during a Crisis: Evidence from Indonesia. Mimeo.

Thomas, D., Frankenberg, E., Beegle, K., & Teruel, G. (2000b). Household Budgets, Household Composition and the Crisis in Indonesia: Evidence from Longitudinal Household Survey Data. Mimeo, RAND.

Thomas, D., Smith, J. P., & Frankenberg, E. (2000c). Lost But Not Forgotten: Attrition in the Indonesia Family Life Survey. *Journal of Human Resources, 36*(3), 556–592.

World Bank. (1998). Indonesia in Crisis: A Macroeconomic Update. Washington DC, July.

Growth and Recession: Underemployment and the Labour Market in the North of England

Anthony Rafferty, James Rees, Marianne Sensier, and Alan Harding

1 Introduction

The North of England constitutes a region of England with a long industrial history in coal mining, steel making, engineering, railways, textiles and shipbuilding, which made it, arguably, the premier manufacturing region in Europe in the middle of the nineteenth century. International competition, first from Europe and America and since the Second World War from Asia, has meant the decline of these staple industries and the

A. Rafferty
Work & Equalities Institute (WEI), Alliance Manchester Business School, University of Manchester, Manchester, UK

J. Rees
Institute for Community Research and Development, University of Wolverhampton, Wolverhampton, UK

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_23

shift to the service sector as an employer of labour. The North of England finds itself now on the periphery of Europe, outside the dynamic core regions contained within the 'pentagon' that has London-Hamburg-Munich-Milan-Paris as its apex cities. However, there are economic success stories for some North of England city regions. At the same time, the recent economic crisis and effects of austerity and public sector cuts provide new challenges to the economy and labour market, increasing levels of under-utilisation of the workforce in the region.

The main aim of this chapter is to examine trends in economic performance and the labour market in the context of the labour market impact of the 2008/2009 recession. We also consider the particular challenges that the North[1] faces as a result of public expenditure constraints. This task is approached in sequential stages that involve analysing a map of GVA change during the long, pre-crisis boom before estimating the pattern of change in the UK labour market through the post-crisis recession and onward to the early stages of economic recovery.[2] A focus is placed on exploring trends in unemployment, involuntary part-time work and time-related underemployment and skill under-utilisation, using over-education as a proxy, specifically through drawing contrasts between the northern regions and London. Time-related underemployment exists when the hours of work of an employed person are insufficient in relation to an alternative employment situation in which the person is willing and available to engage (ILO, 1998). This includes people who are willing to work extra hours in their present job, in an additional job or in a new job in replacement of their current employment. Some estimates suggest that

M. Sensier (✉)
Alliance Manchester Business School, The University of Manchester, Manchester, UK
e-mail: marianne.sensier@manchester.ac.uk

A. Harding
Greater Manchester Combined Authority, Manchester, UK

[1] In this chapter the North refers to the regions of the North West, North East and Yorkshire and Humberside.

[2] These phases refer to the following years: the long boom stretches from 1996 to 2006, the financial crisis and recession is associated with 2007 to 2009 and the recovery period to 2010 and 2011.

more than half of all workers in the UK have experienced a pay cut, a reduction in hours or a loss of benefits since the recession began (Keep Britain Working, 2009). Over-education refers to where people hold qualifications in excess of those required for their job, which may partly reflect skill under-utilisation (Rafferty, 2012).

The plan of this chapter is as follows. The second section discusses key transformations in the economy of the North and what happened during the boom. The third section describes our methods of inquiry and documents the estimated evidence of the North's experience of recession analysing developments across the aforementioned labour market indicators. The fourth section analyses the effect of the government's austerity measures on public sector employment in the North. The final section debates the implications of trends and spatial variation in those trends for future policy.

2 The North During the Boom

The changing economic geography of the North during the last decade of the boom years, seen within the broader national context, is summarised in Rafferty et al. (2013), which maps deflated data (at constant 2000 prices) in % growth of Gross Value Added (GVA) at the NUTS[3] 3 scale between 1996[4] and 2006—the last data point before the onset of the financial crisis in 2007. We also present a table of % growth of GVA in the NUTS 2 regions of the North as these are the units of analysis we focus on in our subsequent labour market analysis.

Economic activity in the UK is dominated by the 'London superregion' which takes in the capital itself and fans out along the major arterial routes that connect it to the rest of the South East, the southern half of the East of England region and into those areas of the South West and Midlands that are best connected to the capital. Beyond this area, the largest GVA concentrations were found in and around the principal provincial metropolitan areas, especially Manchester South, Leeds, York and

[3] NUTS = Nomenclature d'Unités Territoriales Statistiques (Nomenclature of Territorial Units for Statistics). See for details: http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction.

[4] The recession in UK employment is dated to have ended in March 1993 by Sensier and Artis (2011) but the lowest levels of employment are found to be later for Northern regions: November 1994 for the North East, October 1995 for the North West (Artis & Sensier, 2010) and May 1994 for Yorkshire and Humberside.

Newcastle and in other, less continuously urbanised areas in which GVA-rich industrial strengths have traditionally (e.g. in East Merseyside) or more recently (e.g. around Derby with Rolls Royce) been clustered. More modest 'echoes' of the level of GVA growth experienced in the south occurred along the M6, which connects the West Midlands to the North West, and the M62 corridor area which connects Liverpool, in the west, through northern Cheshire and southern Greater Manchester, to Leeds. Similar patterns occurred in and around the principal non-English conurbations, centred on Belfast, Cardiff, Edinburgh and Glasgow. The contrast between the North and the London super-region is not only about the scale of change, however. It also reflects differences in the extent of internal variation. The North, for example, contained two of the seven NUTS 3 areas that experienced GVA decline in the period (Greater Manchester North and West Cumbria) and growth in eastern and northern areas of the North—with the exception of Tyneside—was far more modest than in the 'south of the North'. GVA discrepancies in the London super-region, by contrast, are less stark and are limited to variations within the higher growth bands.

It is important to note two things about these GVA figures. First, they are workplace based and therefore tell us nothing, in and of themselves, about the extent to which residents of NUTS 3 areas benefit from, or are disadvantaged by, patterns of spatial economic concentration and change within 'their' areas. The links between place-specific jobs and those employed within them depend upon travel-to-work patterns and the way in which housing markets 'sort' people of different levels of education and skills into distinct residential areas.

The changes in the period considered here continued a trend that has persisted since the mid-1970s whereby spatial differences in economic performance widened consistently within the UK. This contrasts, domestically, with the 30-year period after Second World War, when disparities narrowed and, internationally, with the experiences of other developed economies in which recent sub-national differences in growth rates have either diverged more slowly (e.g. in the US) or continued to narrow, albeit at a slower rate than previously (e.g. in France, Germany and Italy) (BIS, 2010: 37–38). The explanation for the UK trends that has appeared in official reports—that they are related to 'international economic trends such as globalisation and technological progress' (BIS/CLG, 2010: 7–8)—has not, as yet, been accompanied by evidence on why the

international factors that ostensibly drive sub-national economic divergence operate differently in the UK context.

These overall patterns of change in the North, relative to the rest of the country, during the later 'boom' years are captured by various measures that are sometimes held to be 'drivers' of differential economic performance but are probably best seen as symptoms rather than causes of change. One element of the difference is found in the sector structure of employment. As Green (2010) shows, relative to the UK, all three Northern regions, and especially the North East, remained over-represented in terms of their sectoral share of employment in manufacturing, public administration and health and under-represented in banking, finance and insurance. The differences in sub-Northern economic trajectories during a period of sustained national economic growth were highlighted by Overman (2010), who showed, for example, that GVA per head in Blackpool was half of that in London.

The second geographical message that comes through strongly is the flipside: the relative decline of areas on the periphery (especially port cities such as Liverpool and Hull) and of areas that were once at the heart of the industrial revolution and centres of manufacturing. Overall, the sorts of analyses carried out by economic think tank Centre for Cities reinforce the picture of decline in areas of traditional manufacturing and especially those that are peripheral, have poor environments, outdated amenities, high deprivation and low skills (Webber & Swinney, 2010). There is strong evidence to suggest these trends will remain in place.

A further reason to suggest that we will see an intensification of these trends is what we know from the recent work in the New Economic Geography and in particular that on agglomeration (Fujita & Thisse, 2002). Agglomeration describes the concentration of people and businesses within a geographical space, usually but not always a city-region. In an increasingly knowledge-based economy urban areas with dense concentrations of people and businesses generate 'urbanisation economies', whereby advantages are gained by households as well as firms, regardless of sector, from dense concentrations of economic activity and co-location of workers. Urbanisation economies are partially based on 'economies of scope' which offer agents located in densely populated markets the opportunity to take advantage of positive externalities, such as those associated with knowledge spillovers and improved firm-worker matching—and are particularly associated with the service sector. In practice, it is difficult to disentangle the contribution of these various effects (BIS, 2010).

Recent research has measured agglomeration affects, producing a range of estimates of 2 to 20% increase in productivity from a doubling of economic mass (Brülhart & Mathys, 2008; Sensier et al., 2011). An earlier study of agglomeration in the UK found that a doubling of the working age population in an area is associated with a 3.5% increase in productivity in the area (Rice et al., 2006). It is important to recognise that this measure of 'economic mass' does not necessarily refer to total population but to the importance of proximity and linkages, highlighting the importance of travel to work times, mobility and good infrastructure connectivity (Harding & Rees, 2010). Indeed, another indicator that size does not automatically confer agglomeration benefits is that England's large 'second-tier' cities appear to 'underperform' given the size of their urban populations (Overman & Rice, 2008). This was highlighted in the more detailed research on Greater Manchester in the Manchester Independent Economic Review (MIER, 2009). Other recent research found that agglomeration effects have become more important over time across Europe (Harding et al., 2010).

Differentials in economic/labour market performance would have been even greater without the compensation of the growth in public sector employment over a similar period. In terms of the impact on the North, this growth was a stealthier approach to altering the geographical distribution of employment and economic development, and included both public sector relocation (BBC to Salford, DWP/HMRC civil service cluster in Sheffield) and the general expansion of public sector employment.

This increase in public sector employment was particularly rapid in the 'Labour decade'. Between 1998–1999 and 2008–2009 the share of public sector employment and spending expanded as a percentage of the economy as a whole (Larkin, 2009). Using the Annual Business Inquiry's definition of 'public administration, education and health', public employment grew on average by 2.5% per year between 1998 and 2007, with expansion especially marked in health and education. In terms of urban areas and particularly those in the north it is apparent that central and local governments were to an extent pursuing a 'public sector growth model' (Larkin, 2009). For example, overall 840,000 (69%) of the 1.2 million net additional jobs created in UK cities between 1998 and 2007 were in public administration, education and health. In Manchester, despite its success in replacing manufacturing jobs in the period, witnessing net private sector jobs growth (MIER, 2009), public sector jobs accounted for 59% of net additional jobs. The figure in Leeds was 55% (Larkin, 2009).

One final consideration is that a substantial proportion of the private sector jobs which were created during the period of national economic growth, both in England generally and across the North more specifically, were dependent on the growth in public spending. Some estimates suggest that over half of all private sector jobs created in the last decade were dependent, indirectly, on high levels of public spending (Buchanan et al., 2009). The growth of this parastate activity included the rapid growth of the construction sector as a result of city centre-based 'Urban Renaissance', hospital and school reconstruction, and consultancy and professional services as a result of central and local government contracting.

3 The North's Experience of Recession

Following the recent international financial crisis, UK output first began to contract in the second quarter of 2008, resulting in around a 6% drop in GDP by the end of 2008. Although the drop in GDP witnessed was the deepest and most prolonged for almost 30 years, with increases in unemployment and a slowing of wage growth impacting heavily on the economy, its effect on jobs did not follow the patterns of the previous two recessions, being characterised by a smaller fall than anticipated. The effects of government and Bank of England intervention cushioned impacts on unemployment, through monetary measures (interest rate policy, quantitative easing), demand support measures and public and private labour market support measures (expansion of public sector employment, short time working). Nevertheless, there is an absence of rigorous evaluation of the combined effects of these actions, the longer term consequences, or likely labour market impact once levels of public sector demand are reduced. Furthermore, an alternative possible explanation of the lower-than-expected drop in unemployment was that firms were yet to fully adjust to the economic crisis (Dolphin, 2009), meaning that continued difficult economic conditions are likely to lead to further increases in unemployment.

Table 1 shows the UK regional differences by Government Office Regions (NUTS 1 level), using March 2008 as a baseline. Here growth in unemployment as measured by the Job Seeker's Allowance (JSA) claimant count was most pronounced in Northern Ireland, the North East, the West Midlands and Yorkshire and Humber. In England, outside of the West Midlands, most of the areas that had the biggest increase in unemployment after March 2008 were northern city regions, with areas with

Table 1 Regional unemployment rates based on JSA claimant counts, United Kingdom, March 2008 and September 2009

Region	*March 2008 (%)*	*Sept 2009 (%)*	*Change in %*
Northern Ireland	2.7	6.2	3.5
North East	*3.9*	*7.3*	*3.4*
West Midlands	3.4	6.7	3.3
Yorkshire and Humber	*2.9*	*6.0*	*3.1*
Wales	2.8	5.8	3.0
North West	*3.0*	*5.8*	*2.8*
East Midlands	2.4	5.1	2.7
East	1.9	4.3	2.4
Scotland	2.5	4.7	2.2
South West	1.3	3.5	2.2
South East	1.4	3.5	2.1
London	2.6	4.7	2.1

Source: ONS, NOMIS, authors' calculations

JSA = Job Seeker's allowance

The three regions of the North of England are highlighted in italics

% = 100 × (JSA Count/Economically Active Persons)

high levels of unemployment prior to recession generally being the hardest hit (see Dolphin, 2009).

Table 2 examines the change in JSA claimant counts by NUT3 geographical areas in England between 2008 and January 2011. This confirms broader trends, with areas in the North East ranking high in terms of percentage point increases. Seven out of the top ten ranking areas for percentage point increases in JSA claimant counts in England over this time period are in Northern England.

Although the percentage drop in the employment rate was less substantial than in previous recessions, the percentage drop in the total number of hours worked in the economy was fairly similar (see ONS, 2009, Fig. 2.5). This suggests that, compared to previous recessions, more employers were able to adjust to reduced demand through reductions in working hours rather than the open shedding of labour. Increases in involuntary part-time employment and time-related underemployment support this view (see Walling & Clancy, 2010). In the subsequent analysis, a repeated cross-sectional micro-dataset formed from the spring quarters of the 2007–2011 Special License UK Labour Force Survey (UKLFS) is used to provide insights into this key feature of the recent downturn—the

Table 2 Change in JSA claimant unemployment rate by NUTS 3 regions in the North of England, March 2008 to January 2011

NUTS 3 region	*JSA rate Jan 2011*	*Change in JSA rate (March 2008 to Jan 2011)*	*% change in JSA rate*	*Rank of change in JSA rate in England (out of 99)*
Kingston upon Hull, City of	7.7	3.2	69.7	1
South Teesside	7.1	2.9	68.6	2
North and North East Lincolnshire	5.4	2.6	88.7	4
Hartlepool and Stockton-on-Tees	6.0	2.6	74.2	5
Blackpool	6.1	2.6	72.2	7
Barnsley, Doncaster and Rotherham	4.9	2.3	91.8	8
Darlington	5.0	2.2	75.4	9
Sheffield	4.4	2.1	90.5	12
Calderdale, Kirklees and Wakefield	4.4	2.1	89.4	13
Bradford	5.0	2.1	71.5	14
Sunderland	5.2	2.0	61.1	17
Sefton	5.0	2.0	63.0	18
Leeds	4.3	2.0	83.5	19
Derby	4.5	2.0	80.1	20
Greater Manchester North	4.6	2.0	75.8	23
Durham CC	4.0	1.8	86.2	32
Tyneside	5.0	1.8	54.2	33
Halton and Warrington	4.1	1.8	77.4	34
Greater Manchester South	4.3	1.8	74.7	35
East Merseyside	5.3	1.8	52.8	36
Liverpool	6.7	1.7	33.7	38
Blackburn with Darwen	4.3	1.7	67.2	41
East Riding of Yorkshire	3.1	1.5	93.3	47
Northumberland	3.9	1.4	55.1	59
West Cumbria	3.5	1.4	61.6	60
York	2.7	1.3	87.6	63
North Yorkshire CC	2.6	1.3	100.4	64
Cheshire CC	2.8	1.2	81.8	74
Wirral	4.5	1.1	33.0	75
Lancashire CC	2.9	1.1	64.2	83
East Cumbria	2.0	0.9	76.9	93

Source: ONS NOMIS, own calculations

CC = County Council

Unemployment rate = 100 × (JSA Count/Estimate of Economically Active Population)

% change in JSA rate = 100 × ((JSA rate Jan 2011 - JSA rate Mar 2008)/JSA rate Mar 2008)

lower-than-expected rise in unemployment, compared to previous recession. Patterns of time-related underemployment (see Walling & Clancy, 2010), involuntary part-time work, over-education and the extent to which 'graduates' holding NVQ level 4+ qualifications work in non-graduate occupations, as defined by the SOC(HE) classification of graduates jobs (Elias & Purcell, 2004), are examined.

During recession employers can use reductions in working hours as a way of maintaining staff to avoid shedding labour and the recruitment and training costs associated with subsequent hiring. Such practices may occur through voluntary arrangements, but also through reducing working hours offered to new recruits, increasing levels of involuntary part-time work or broader time-related underemployment. One potential consequence of this is that the extent of job creation witnessed for a corresponding unit of GDP growth during recovery is less than when increases in labour demand are first met by employers increasing working hours amongst retained but time underemployed staff rather than through open recruitment or job creation.

Time-related underemployment and involuntary part-time work were particularly marked in certain Northern NUTS 2 areas. Using the UK Labour Force Survey (LFS), Table 3 indicates that in 2009/2010, seven out of eleven Northern NUTS 2 areas were ranked in the top ten in England (out of 30) for levels of involuntary part-time work. The fourth column of the table provides estimates of percentage change. Given this change is estimated based on a fraction of the overall area sub-samples the figures present should be treated as fairly rough estimates of sub-regional differences. Across the UK, NUTS 2 areas experienced around a 50% increase in the proportion of part-time work which was involuntary between 2007/2008 and 2009/2010. Some of the highest rates of growth in involuntary part-time employment appeared to be in East Yorkshire and North Lincolnshire, South Yorkshire and Greater Manchester. Growth appears to have been lower in inner London, although starting from a considerably higher baseline in 2007/2008. The growth in many other areas could therefore reflect a process of catching up to inner London levels.

Levels of broader time-related underemployment in 2009/2010 were fairly similar across the UK counties (Table 4). Three outliers in the Northern regions however were East Yorkshire and North Lincolnshire, Greater Manchester and Tees Valley which were ranked 1st, 4th and 6th in England respectively. Although Inner London was ranked number 1 for

Table 3 Change in involuntary part-time work, NUTS 2 regions, North of England, 2007/2008 to 2009/2010

NUTS2 region	*% employed part-time, 2009/2010*	*% of part-time involuntary*	*Rank[a] (NUTS2 England)*	*% change involuntary part-time (base 2007/2008)*	*Sample[b] number*
East Yorkshire and North Lincolnshire	29.0	19.4	4	+75.6	1316
South Yorkshire	28.2	16.1	10	+71.6	1808
Greater Manchester	24.4	18.3	7	+61.7	3466
Tees Valley and Durham	25.2	19.2	5	+59.8	1611
Merseyside	24.8	16.2	9	+58.4	1714
North Yorkshire	28.0	12.8	22	+52.7	1117
Cheshire	23.0	10.5	27	+52.1	1541
Lancashire	24.2	16.3	8	+49.4	2027
Northumberland	23.8	19.7	2	+45.0	2097
West Yorkshire	24.4	13.1	20	+25.5	3221
Cumbria	26.6	15.5	11	+17.4	736
Inner London	17.2	20.0	1	+6.1	2800
Rest of England	24.9	13.7		+50.0	44,820

Source: Special License UK LFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

[a] Rank is ordinal position in English NUTS 2 areas (out of 30) based on % of part-time involuntary

[b] Number of observation is un-weighted base

involuntary part-time employment, within the multivariate analysis using logistic regression to control for compositional differences in the demographic characteristics of individuals living in different areas, Table 5 indicates that the risk of involuntary part-time work was significantly higher in the Tees Valley and Durham (Odds Ratio (OR) = 1.27, $p < 0.05$), Northumberland (OR = 1.34, $p < 0.01$), and East Yorkshire and North Lincolnshire (OR=1.35, $p < 0.05$). Individual differences were apparent with the odds of experiencing involuntary part-time employment [Model I] and broader time-related underemployment [Model II] being higher for women, people with lower educational attainment and those with shorter job tenure.

In addition to time-related underemployment, a further potential manifestation of underemployment is that due to a lack of suitable employment, people take on paid work below their skills or qualification levels.

Table 4 Time-related underemployment, NUTS2 regions, North of England, 2007/2008 to 2009/2010

NUTS2 region	*% under-employed of employed 2009/2010*	*Rank[a] (NUTS2 England)*	*% change (base = 2007/2008)*	*Sample number[b]*
East Yorkshire and North Lincolnshire	12.0	1	+85.2	1316
Greater Manchester	11.0	4	+56.8	3466
South Yorkshire	10.7	8	+53.3	1808
West Yorkshire	9.9	17	+46.2	3221
Merseyside	9.4	24	+40.1	1714
Tees Valley and Durham	10.9	6	+33.9	1611
Lancashire	10.3	15	+27.9	2027
North Yorkshire	10.8	8	+26.3	1117
Cheshire	8.0	29	+21.4	1541
Northumberland	9.6	21	+19.2	2097
Cumbria	8.1	27	-8.5	736
Inner London	9.6	22	+28.4	2800
Rest of England	9.9		+39.0	44,820

Source: Source: Special License UK LFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

[a] Rank is ordinal position in English NUTS 2 areas (out of 30) based on % underemployed of employed

[b] Number of observation is unweighted base

Skills-related underemployment refers to the involuntary confinement to employment that does not sufficiently utilise skill levels or qualifications due to a lack of better matched opportunities (ILO, 1998). Skills-related underemployment is important in that it represents an additional dimension of macro-economic under-productivity and mismatch in the labour market. At a regional level, a lack of employment opportunities that meets skills or qualification levels may further lead to internal economic migration or 'brain drain' to other geographical areas with greater opportunities, particularly amongst the younger or more highly qualified.

Although skills-related underemployment is difficult to measure in survey data, measures of over-education provide one potential proxy. Over-education refers to where people hold qualifications that are in excess of those required for their job (e.g. Rumberger, 1987; Rafferty, 2012). In comparison to those better matched to their job, research indicates that

Table 5 Logistic regression results for involuntary part-time work and time-related underemployment, NUTS2 regions, UK 2007/2008 and 2009/2010

Variables	*Model I: involuntary part-time employment*		*Model II: time-related underemployment*	
	Odds ratio	*SE*	*Odds ratio*	*SE*
Region dummies (0 = inner London)				
Cumbria	1.25	0.20	0.99	0.10
South Yorkshire	1.11	0.13	0.99	0.07
Greater Manchester	1.08	0.11	1.00	0.06
Merseyside	1.06	0.13	0.91	0.07
Lancashire	1.03	0.12	1.02	0.07
North Yorkshire	0.95	0.14	1.13	0.10
West Yorkshire	0.92	0.10	0.91	0.06
Cheshire	0.67*	0.10	0.81*	0.07
Tees Valley and Durham	1.27*	0.15	1.10	0.08
Northumberland	1.34**	0.15	1.03	0.07
East Yorkshire and North Lincolnshire	1.35*	0.17	1.03	0.09
Rest of UK	0.94	0.07	0.94	0.05
Other variables				
Years of schooling	0.97**	0.00	0.94**	0.00
Job tenure	0.99**	0.00	1.00**	0.00
Female (0 = male)	1.61**	0.05	1.18**	0.02
Age	1.01**	0.00	0.99**	0.00
Year 2009/2010 (0 = 2007/2008)	1.59**	0.05	1.47**	0.03

Source: Special License UK LFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

SE = Standard Error

Significance levels: * = $p < 0.05$ ** = $p < 0.01$

over-educated people exhibit wage penalties (Hartog, 2000), lower job satisfaction (Battu et al., 1999) and higher turnover (Sloane et al., 1999).

A modal 'realised matched' measure of over-education was used to explore patterns of over-education amongst the general population (see Hartog, 2000; Verdugo & Verdugo, 1989), whereas the SOC(HE) classification of graduate and non-graduate jobs was used to examine 'graduate-level' over-education amongst people holding degree-level (NVQ level 4/5) qualifications (see Elias & Purcell, 2004). For the first measure, the modal NVQ level qualification (levels 1 to 4/5) for each

Table 6 Over-education, NUTS2 regions, North of England, 2009/2010 and change from 2007/2008

NUTS2 region	*% over-educated*[a]	*Rank (NUTS2 England)*	*% change (base = 2007/2008)*	*Sample*[b] *number*
East Yorkshire and North Lincolnshire	28.1	3	+6.3	1316
Cumbria	27.2	6	-3.7	736
Lancashire	26.9	7	+14.2	2027
Greater Manchester	26.7	9	+1.2	3466
Merseyside	26.7	10	-0.2	1714
Tees Valley and Durham	26.6	11	-2.7	1611
Northumberland	26.5	13	+1.5	2097
North Yorkshire	25.1	16	-2.9	1117
South Yorkshire	25.0	17	+1.7	1808
Cheshire	24.5	18	+2.7	1541
West Yorkshire	23.7	20	-5.5	3221
Inner London	19.8	30	-0.6	2800
Rest of England	24.3	–	+6.5	44,820

Source: Special License UKLFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

[a] Modal measure. Respondent is defined as over-educated if their level of educational attainment exceeds the modal NVQ level for their SOC 3-digit occupation

[b] Number of observation is unweighted base

3-digit Standard Occupational Classification (SOC 2000) code was calculated. This was compared to the highest qualification level held by respondents. Respondents were defined as over-educated if their qualification levels were above the modal level for their occupation. Northern NUTS 2 areas contributed six out of ten of the highest ranking areas in England for both overall and graduate (SOC(HE)) levels of over-education (Tables 6 and 7). The lowest levels in contrast were witnessed in Inner London. Four out of five of the areas with the highest levels of graduate level over-education were in Northern England (East Yorkshire and Lincolnshire, Greater Manchester, Lancashire and Merseyside).

Compared to patterns of involuntary part-time work and time-related underemployment, changes between 2007/2008 and 2009/2010 were less marked for overall levels of over-education, although again the smaller sub-samples used to derive percentage change figures make these figures fairly rough estimates at the sub-regional level (Table 7). One potential

Table 7 Percentage of employed NVQ level 4/5 holders working in SOC(HE) 'Non-Graduate' occupations, NUTS2 regions, North of England, 2009/2010 and change from 2007/2008

NUTS2 regions	*% in non-graduate occupations*	*Rank[a] (NUTS2 England)*	*% change (base = 2007/2008)*	*Sample number[b]*
East Yorkshire and North Lincolnshire	30.5	2	–1.3	399
Greater Manchester	28.1	3	–4.1	1231
Lancashire	28.0	4	27.6	718
Merseyside	27.8	5	–6.3	583
Northumberland	27.1	7	13.2	657
Tees Valley and Durham	26.6	8	15.7	494
Cumbria	24.7	16	–17.1	240
South Yorkshire	22.2	20	–4.2	538
Cheshire	21.2	23	-8.2	546
West Yorkshire	20.5	26	–16.5	1085
North Yorkshire	20.0	29	–31.5	431
Inner London	17.8	30	0.3	1600
Rest of England	23.5		5.3	15,692

Source: Special License UKLFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

[a] Rank is ordinal position in English NUTS 2 regions (out of 30)

[b] Number of observation is unweighted base. SOC (HE) definition of graduate occupations based on expert ranking, detailed occupational descriptions and statistical information on qualification concentrations within occupations (see Elias & Purcell, 2004)

explanation is that lower skilled jobs which over-educated people are more greatly represented within are more likely to be destroyed during economic downturn meaning an apparent improvement in levels of over-education is in actuality explained by the over-educated being more likely to become unemployed. It may also be the case that whereas time-related underemployment is an early recession effect, patterns of skill-underutilisation do not manifest until later on as long-term employment increases and people reduce their labour market expectations or reservation wages to escape unemployment, or they may be pushed towards any type of work due to JSA eligibility criteria or means-testing after 6 months of claiming JSA.

Table 7 goes on to consider the extent to which graduates (holding NVQ level 4+ qualifications) were likely to work in graduate or non-graduates jobs as defined by the SOC (HE) classification (Elias & Purcell, 2004). Inner London experienced the lowest levels of 'graduate over-education' whereas some of the highest levels were found in East Yorkshire and North Lincolnshire, and Greater Manchester. Logistic regression (Table 8) confirmed that the Northern areas (and the rest of the UK) experienced a relative penalty in terms of overall and graduate over-education compared to Inner London after controlling for education levels and other demographic factors.

Table 8 Logistic regression results for over-education and representation of graduates in non-graduate jobs. NUTS 2 regions, UK 2007/2008 and 2009/2010

Variables	*Model I: over-education (modal)*		*Model II: non-graduate job*	
	Odds ratio	*SE*	*Odds ratio*	*SE*
Region dummies (0 = inner London)				
Tees Valley and Durham	1.43**	0.07	1.32**	0.12
Northumberland	1.39**	0.07	1.37**	0.11
Cumbria	1.53**	0.10	1.50**	0.18
Cheshire	1.22**	0.07	1.28**	0.11
Greater Manchester	1.36**	0.06	1.58**	0.11
Lancashire	1.30**	0.06	1.36**	0.11
Merseyside	1.40**	0.07	1.62**	0.14
East Yorkshire and North Lincolnshire	1.42**	0.08	1.80**	0.17
North Yorkshire	1.30**	0.08	1.40**	0.14
South Yorkshire	1.25**	0.06	1.18**	0.11
West Yorkshire	1.20**	0.05	1.20*	0.09
Rest of UK	1.18**	0.04	1.26**	0.06
Other variables				
Years of schooling	0.98**	0.01	0.88**	0.00
Job tenure	1.00*	0.01	1.00**	0.00
Female (0 = male)	0.93**	0.01	0.93**	0.02
Age	0.99**	0.00	0.97**	0.00
Year 2009/2010 (0 = 2007/2008)	1.07**	0.01	1.04	0.02

Source: Special License UKLFS Micro-datasets, own calculations (men 16–64 years and women 16–59 years): seasonally unadjusted; weighted to 2009 mid-year population estimates

SE = Standard Error

Significance levels: * = $p < 0.05$ ** = $p < 0.01$

Regional differences however only tell part of the story of the labour market impact of the recession, meaning that some groups of workers have experienced a greater impact of recession across regions. For example, the 2008/2009 recession hit younger men and women harder, as evidenced by the rise in youth unemployment (Rubery & Rafferty, 2013). In addition to higher unemployment, Table 9 further indicates that levels of time-related underemployment and over-education are also greater for

Table 9 Youth underemployment and over-education by NUTS 2 region

	<25 yrs	*26–45 yrs*	*46+ yrs*	*<25 yrs*	*26–45 yrs*	*46+ yrs*
	Involuntary part time			**Time-related underemployment**		
Tees Valley and Durham	8.1	3.6	4.7	14.6	9.7	10.4
Northumberland	8.1	3.8	4.2	14.8	9.9	7.1
Cumbria	8.6	3.6	2.7	20.3	6.7	4.4
Cheshire	6.0	1.5	1.8	13.8	7.2	5.9
Greater Manchester	8.9	3.4	3.2	17.0	10.3	8.4
Lancashire	8.8	2.3	3.8	20.6	8.7	6.8
Merseyside	6.9	3.2	3.8	17.5	7.8	7.7
East Yorkshire and North Lincolnshire	9.9	5.2	3.6	15.7	12.6	8.8
North Yorkshire	6.6	3.2	2.1	20.8	9.2	6.9
South Yorkshire	5.1	4.8	3.8	13.0	13.0	6.7
West Yorkshire	5.5	2.5	2.8	16.2	9.5	6.7
Inner London	4.3	2.8	4.4	15.2	8.1	10.4
Rest of England	7.0	2.5	2.9	16.2	9.3	7.5
	Over-educated (modal)			**% graduates in SOC(HE) non-graduate jobs**		
Tees Valley and Durham	37.6	24.5	23.0	38.4	27.2	19.6
Northumberland	39.9	23.9	23.8	56.1	23.4	22.2
Cumbria	35.9	24.5	26.6	37.1	19.6	27.9
Cheshire	39.3	22.2	19.9	49.1	18.8	17.2
Greater Manchester	40.9	23.1	23.2	56.3	20.6	27.7
Lancashire	36.1	25.2	24.2	52.8	25.4	23.2
Merseyside	47.0	22.5	23.0	62.9	20.7	25.7
East Yorkshire and North Lincolnshire	40.2	27.6	21.1	63.9	29.1	22.6
North Yorkshire	34.1	22.3	23.7	33.0	18.4	18.7
South Yorkshire	35.6	23.3	22.2	48.5	21.1	16.8
West Yorkshire	38.7	20.8	19.0	41.7	17.9	16.8
Inner London	31.5	17.6	18.5	50.0	21.3	19.2
Rest of England	39.7	21.9	20.1	31.5	15.6	17.9

people under 25 years of age, although this may partly relate to lower levels of employment experience (Rafferty, 2012). Notably, in relation to graduate-level (SOC (HE)) over-education, levels of over-education among younger employees were particularly high in Greater Manchester, Merseyside, East Yorkshire and North Lincolnshire. In the latter two NUTS 2 areas, over six out of ten employed graduates under the age of 25 worked in non-graduate occupations compared to less than three out of ten in the older age categories. Although over time greater employment experience or between-job transitions may help people move into employment better suited to their qualification levels, the extent of youth over-education remains comparatively high.

Austerity and Public Sector Employment

Towards the end of the long boom, there were warnings from commentators that the North had become 'dependent' on public sector jobs growth. The North would therefore be particularly vulnerable to any shrinkage in employment resulting from public spending cuts to address the structural deficit. It follows that there will be impacts on the wider economy of the North and this has raised concerns about which places might be relatively vulnerable or resilient. Growth in the public sector was a key driver of overall jobs growth over the last decade, with the share of public sector employment and spending both expanding as a percentage of the economy as a whole. It has already been noted that over two thirds of the net additional jobs created in cities in the UK between 1998 and 2007 were in the public sector. From a northern perspective, as Dolphin (2009) notes, public spending as a share of GVA in 2007/2008 was 52% in the North East, 47% in the North West and 44% in Yorkshire and the Humber, as against 38% for the UK as a whole.

Because public spending has been higher in the North, many northern places, particularly those characterised by high concentrations of deprivation, are being and will be disproportionately affected by public expenditure cuts. Table 10 lists the local authority areas that are most and least affected by local spending cuts. It shows clearly that it is local authority areas in the North, along with the poorest Inner London boroughs, which face the severest cuts. By contrast, the areas experiencing the lowest spending reductions, are, with only two exceptions (Warwickshire and Cheshire East) found in the South East. As Wilks-Heeg (2011) has found, drawing upon redundancy plans submitted to government, the effect of cuts is to

Table 10 Planned revenue spending power changes by local authority area: most and least affected areas, England, 2011/2012

Local authority	*Change in estimated 'revenue spending power' 2011/2012[a] (£m)*	*Change in estimated 'revenue spending power' 2011/2012 (%)*	*Transition grant (£m)*	*Change in estimated 'revenue spending power' 2011/2012—post transition grant (%)*
Biggest % losers, ranked by £m spending reduction				
Liverpool	-72.2	-11.34	15.550	-8.90
Manchester	-68.9	-11.04	13.332	-8.90
Hackney	-38.7	-10.46	5.776	-8.90
Newham	-38.5	-10.46	5.731	-8.90
Tower Hamlets	-37.2	-9.90	3.767	-8.90
Doncaster	-29.2	-9.60	2.118	-8.90
Kingston upon Hull	-28.4	-9.72	2.391	-8.90
Knowsley	-23.4	-10.72	3.968	-8.90
Rochdale	-21.7	-8.91	0.031	-8.90
South Tyneside	-19.8	-10.49	3.009	-8.90
Blackburn with Darwen	-18.1	-10.54	2.806	-8.90
St Helens	-18.1	-9.39	0.935	-8.90
Middlesbrough	-15.4	-9.12	0.364	-8.90
North East Lincolnshire	-15.3	-9.14	0.396	-8.90
Halton	-12.6	-8.92	0.026	-8.90
Hartlepool	-12.0	-10.33	1.661	-8.90
Lowest % change, ranked by £m spending reduction				
Essex	-13.0	-1.31	0.000	-1.31
Hertfordshire	-12.5	-1.55	0.000	-1.55
Norfolk	-11.7	-1.67	0.000	-1.67
Hampshire	-7.8	-0.95	0.000	-0.95
East Sussex	-6.3	-1.54	0.000	-1.54
Warwickshire	-5.1	-1.32	0.000	-1.32
Cheshire East	-4.6	-1.63	0.000	-1.63
West Sussex	-3.7	-0.65	0.000	-0.65
Surrey	-2.5	-0.31	0.000	-0.031
Buckinghamshire	-2.0	-0.60	0.000	-0.60
Poole	-1.1	-0.97	0.000	-0.97
Windsor and Maidenhead	-1.1	-1.06	0.000	-1.06
Richmond upon Thames	-1.0	-0.61	0.000	-0.61
Wokingham	-0.7	-0.63	0.000	-0.63
Dorset	0.7	0.25	0.000	0.25

[a]Includes local authority spending plus NHS spending on social care

produce far heavier job losses in the North than in southern regions. Because local authorities are important procurers of services, it must be expected that a similar, disproportionate pattern of job losses amongst private and voluntary sector providers of local goods and services will also occur.

Overall, the patterns described above point in the same direction, that is, urban areas that have been more reliant on public sector expansion in the last decade or so, and those that exhibit higher proportions of public sector employment are more vulnerable to public spending cuts and because of their weaker private sector base are likely to be less resilient to recession (see Martin, 2012). Nationally this is particularly true of the North in general and more so of the traditionally weaker, ex-industrial and marginal parts of the North. Nevertheless, it remains uncertain what the downstream impacts—on supply chains, consumer spending and local economies—will be and hence on longer term patterns of population and economy.

Although the initial period of recession disproportionately hit more 'cyclical' male-dominated industries such as construction and manufacturing, a subsequent increase in job losses within the service sector and the greater representation of women within the public sector in 2011–2012 saw a continuation of job loss among female employees whilst unemployment stabilised among men (Rubery & Rafferty, 2013). Table 11 indicates that overall levels of public sector employment are particularly high in some Northern NUTS 2 areas such as Merseyside (35.4%) and Northumberland (35.6%). However, across all areas women are around 15–24% more likely to be employed in the public sector than men are. Given that the full impact of the spending cuts on job loss are yet to be realised, many of the Northern areas, and women in particular, may be particularly exposed to further public sector job loss.

4 Conclusions

The purpose of this chapter is to provide insights into recent developments in the economy and labour market of the North of England, particularly following the economic crisis that began in 2007/2008, through drawing contrasts with the rest of England and in particular the inner London area. The findings reveal that beyond unemployment, broader labour market problems and recessionary impacts are concealed within employment. In addition to relatively high levels of unemployment, across

Table 11 Percentage of total employment in public sector employment by NUTS 2 Region 2011

	Total		*Men*		*Women*		*Difference (F-M)*
	% public sector	*% change (base = 07/8)*	*% public sector*	*% change (base = 07/8)*	*% public sector*	*% change (base =07/8)*	
Tees Valley and Durham	29.7	0.7	21.2	14.6	38.3	-7.3	17.1
Northumberland	35.6	11.0	25.9	12.1	45.8	9.3	19.9
Cumbria	25.7	–4.4	16.9	–10.7	35.3	–1.8	18.4
Cheshire	23.2	–4.9	14.1	–7.8	33.0	–3.1	18.9
Greater Manchester	29.5	11.8	19.4	15.9	40.3	9.4	20.9
Lancashire	31.8	7.4	20.2	0.5	44.4	11.0	24.2
Merseyside	35.3	8.7	25.8	12.3	45.4	7.7	19.5
East Yorkshire and North Lincolnshire	25.2	–10.8	16.0	–20.3	34.6	–6.9	18.7
North Yorkshire	34.7	9.5	27.2	8.4	42.2	8.9	15.0
South Yorkshire	30.9	8.4	20.4	15.0	41.3	1.6	20.9
West Yorkshire	27.4	5.4	18.3	14.6	37.2	–0.4	18.9
Inner London	25.4	–2.7	19.1	–8.8	33.0	2.1	14.0
Rest of England	27.3	6.1	19.3	7.7	36.1	4.8	16.8

Source: Special License LFS, own calculations

indicators of time-related underemployment and over-education, many of the areas in the North of England experienced comparatively high levels of under-utilisation within paid work. These issues appear particularly acute for younger people. For one thing, levels of non-graduate employment among graduates remain comparatively high in the North of England compared to London and the rest of England generally. Tentatively this may suggest a particular under-utilisation of the skills and talent of the workforce in the Northern regions. The full impact of the employment effects of the economic crisis and on-going expenditure cuts and reductions in public sector employment are unlikely to be fully realised at the time of the current study. Nonetheless, looking towards the future the comparatively high levels of public sector employment in many of the northern regions, particularly among women, create a further potential exposure to job loss and impediment to labour market recovery.

There are some important policy issues arising from this discussion. Much of the evidence we have considered suggests that trends towards the

widening of disparities, between the North and the London super-region, between the functional economic areas of the North and within the labour markets of those functional economic areas, are set to continue. It appears highly unlikely that the policy aspirations of the current Coalition government for spatial rebalancing (HM Government, 2010) are enough to attenuate, far less reverse these trends, particularly in a context in which a 'presumption in favour of development' and government proposals to reward local authorities financially for new commercial and residential development (notably through tax increment financing and a 'new homes bonus'). This, in addition to an overarching commitment to a 'new localism' in which local resources and leadership are expected to play a more important role, will tend to favour areas already favoured by the market.

Even if recovery from the recent recession follows the pattern of previous, less severe downturns, it is likely to take until the end of or beyond the current Parliament in 2015 for national output to return to pre-crisis levels and a further two to three years for employment levels to do the same. There will also be a spatial dimension to recovery whereby Northern output and employment growth will likely lag behind that of Southern regions. Furthermore, irrespective of its exact geography, recovery is likely to drive further polarisation within the labour market, and in household incomes, as manufacturing becomes still less employment-intensive and private service sector employment continues to be characterised by strong income differentials between high and low value activity. In this context, it must be expected that policy initiatives that aim to remove barriers to new development and to place an onus on the mobilisation of local resources will reward those areas that are in the strongest position with respect to market development and are therefore likely to exacerbate the disparities, at various spatial scales as identified.

This is not to argue that explicit policy initiatives that are being targeted upon northern and midland England—including Enterprise Zones and Regional Growth Fund investments—are incapable of mitigating the impact of these trends to some degree. Spatial rebalancing between North and South is also facilitated, to some extent, by natural endowments, as is the case, for example, in relation to renewable energy investments. As we have seen, though, the impact of changes in redistributive forms of implicit policy—especially expenditure cuts and their employment impacts—is working in the opposite direction and reinforcing disparities.

It is not unusual for explicit and implicit spatial policies to counteract one another, so the policy challenges arising from aspirations to rebalance

the economy spatially are by no means new. They are particularly acute, however, with respect to the emphasis increasingly being placed, within policy debates, on the importance of agglomeration and the extent to which 'place blind' implicit policies are seen as relating to the increasing tendency of high-level economic activities to group together in the larger, denser urban areas with the 'thickest' labour markets. At one level, a shift to a more localised approach to economic development appears consistent with the view that spatial economic policy should focus upon managing agglomeration forces more effectively and spreading its benefits in that it should result in choices that enable firms and households to locate in the places they wish to be. It appears inconsistent, however, with certain national investment decisions, particularly in relation to major infrastructure, research and technology and mega-projects (Cross rail for example, and recently the Olympic Games) that continue to be focused disproportionately in the South. This brief discussion highlights the continuing need for clear policy aims with respect to spatial development priorities and for consistency in the way they are aligned and supported across scales of government.

Acknowledgement The authors would like to thank Phil Rees for help drafting this chapter, Alison Manson for producing the map and the two anonymous referees for their comments. We are grateful to the Economic and Social Research Council (grant number RES-000-22-4304) for partly funding this research and to the Office for National Statistics and Economic and Social Data Services for making the UKLFS Special License micro data available.

References

Artis, M. J., & Sensier, M. (2010, February). *Tracking Unemployment in the North West Through Recession and Forecasting Recovery.* New Economy Working Paper Series, No. NEWP: 02. http://neweconomymanchester.com/stories/993-new_economy_working_papers_archive

Battu, H., Belfield, C. R., & Sloane, P. (1999). Over-Education Among Graduates: A Cohort View. *Education Economics, 7*, 21–38.

BIS. (2010). Regional Growth Fund. http://www.bis.gov.uk/RGF

BIS/CLG. (2010). *Understanding Local Growth.* BIS Economics Paper No. 7. Department for Business, Innovation and Skills and Department of Communities and Local Government. BIS.

Brülhart, M., & Mathys, N. A. (2008). Sectoral Agglomeration Economies in a Panel of European Regions. *Regional Science and Urban Economics, 38*, 348–362.

Buchanan, J., Froud, J., Johal, S., Leaver, A., & Williams, K. (2009). *Undisclosed and Unsustainable: Problems of the UK National Business Model*. CRESC Working Paper Series, Working Paper No. 75.

Dolphin, T. (2009). *The Impact of Recession on Northern City-Regions*. Joseph Rowntree Foundation/ippr. http://www.ippr.org/publications/55/1731/the-impact-of-the-recession-on-northern-city-regions

Elias, P., & Purcell, K. (2004). *SOC (HE): A Classification of Occupations for Studying the Graduate Labour Market*. Researching Graduate Careers Seven Years On, Research Paper No. 6. Institute for Employment Research.

Fujita, M. M., & Thisse, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge University Press.

Green, A. E. (2010). Labour Market Trends and Issues: Challenges Facing the North of England, Moving Forward: The Northern Way, Turning Points, Drivers and Assets Papers. Editors: CURDS/SERC. http://www.thenorthernway.co.uk/downloaddoc.asp?id=794

Harding, A., & Rees, J. (2010, March). *Connectivity and the North of England*. Northern Way Turning Points Symposium Working Papers. http://www.thenorthernway.co.uk/downloaddoc.asp?id=795

Harding, A., Rees, J., & Sensier, M. (2010, October). *Agglomeration and governance: Greater Manchester in a European Perspective*. New Economy Working Paper Series, No. NEWP: 05. http://neweconomymanchester.com/stories/993-new_economy_working_papers_archive

Hartog, J. (2000). Overeducation and Earnings: Where Are We, Where Should We Go? *Economics of Education Review, 19*(2), 131–147.

HM Government. (2010). *The Coalition: Our Programme for Government*. Cabinet Office.

ILO. (1998). *Resolution Concerning the Measurement of Underemployment and Inadequate Employment Situations*. Adopted by the Sixteenth International Conference of Labour Statisticians. International Labour Organization.

Keep Britain Working. (2009). Press Release. http://www.reedpressoffice.co.uk/newsArticle/Keep+Britain+Working+campaign+launched/

Larkin, K. (2009). *Public Sector Cities: Trouble Ahead*. Centre for Cities. http://www.centreforcities.org/assets/files/09-07-16%20Public%20sector%20cities.pdf

Martin, R. (2012). Regional Economic Resilience, Hysteresis and Recessionary Shocks. *Journal of Economic Geography, 12*(1), 1–32.

MIER. (2009). *Understanding Labour Markets, Skills and Talent*. Manchester Independent Economic Review. http://www.manchester-review.org.uk/

ONS. (2009, May 14). The Impact of the Recession on the Labour Market. *ONS News Release*. http://www.ons.gov.uk/ons/rel/lmac/impact-of-the-recession-on-the-labour-market/impact-of-the-recession-on-the-labour-market/index.html

Overman. (2010). Election Analysis, Urban Renewal and Regional Growth: Muddled Objectives and Mixed Progress. Centre for Economic Performance. http://cep.lse.ac.uk/pubs/download/ea014.pdf

Overman, H. G., & Rice, P. (2008). *Resurgent Cities and Regional Economic Performance*. SERC Policy Paper 1. http://www.parisschoolofeconomics.eu/IMG/pdf/Overman2-PSE-MEEDM.pdf

Rafferty, A., Rees, J., Sensier, M., et al. (2013). Growth and Recession: Underemployment and the Labour Market in the North of England. *Applied Spatial Analysis and Policy, 6*, 143–163. https://doi.org/10.1007/s12061-013-9089-4

Rafferty, A. L. (2012). Ethnic Penalties in Graduate Level Over-Education, Unemployment and Wages: Evidence from Britain. *Work, Employment & Society, 26*(6), 987–1006.

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial Determinants of Productivity: Analysis for the Regions of Great Britain. *Regional Science and Urban Economics, 36*, 727–752.

Rubery, J., & Rafferty, A. (2013). Women and Recession, Revisited. *Work, Employment, and Society, 27*(3), 414–432.

Rumberger, R. W. (1987). The Impact of Surplus Schooling on Productivity and Earnings. *Journal of Human Resources, 22*(1), 24–50.

Sensier, M., & Artis, M. J. (2011). *Tracking Unemployment in Wales Through Recession and Into Recovery*. Spatial Economics Research Centre Discussion Paper Series, London School of Economics, No. SERCDP0079. http://rlab.lse.ac.uk/_new/publications/abstract.asp?index=3808

Sensier, M., Curran, D., & Artis, M. J. (2011). *Investigating Agglomeration Economies in a Panel of European Cities and Regions*. Spatial Economics Research Centre Discussion Paper Series, London School of Economics, No. SERCDP0078. http://rlab.lse.ac.uk/_new/publications/abstract.asp?index=3807.

Sloane, P. P., Battu, B., & Seaman, P. P. (1999). Overeducation, Undereducation and the British Labour Market. *Applied Economics, 31*, 1437–1453.

Verdugo, R., & Verdugo, N. (1989). The Impact of Surplus Schooling on Earnings. *Journal of Human Resources, 22*(4), 690–695.

Walling, A., & Clancy, G. (2010). Underemployment in the UK Labour Market. *ONS Economic and Labour Market Review, 4*(2), 16–24.

Webber, C., & Swinney, P. (2010). *Private Sector Cities: A New Geography of Opportunity*. Centre for Cities. http://www.centreforcities.org/assets/files/10-06-07%20Private%20Sector%20Cities%20web.pdf

Wilks-Heeg, S. (2011). You Can't Play Politics with People's Jobs and People's Services: Localism and the Politics of Local Government Finance. *Local Economy, 26*(8), 635–651.

The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach

Bruno Arpino, Chiara Pronzato, and Lara Patrício Tavares

1 Introduction

Childcare arrangements are a key element in the work decisions taken by mothers insofar as they are fundamental to the ability of women to juggle childcare and participation in the labour market. In the traditional

B. Arpino
Department of Statistics, Computer Science, Applications of the University of Firenze, Firenze, Italy
e-mail: bruno.arpino@unifi.it

C. Pronzato (✉)
Department of Economics and Statistics, University of Turin, Turin, Italy
e-mail: chiaradaniela.pronzato@unito.it

L. P. Tavares
Instituto Superior de Ciências Sociais e Políticas, Universidade de Lisboa, Lisbon, Portugal
e-mail: lara.tavares@edu.ulisboa.pt

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_24

economic models of female labour supply, where formal childcare is usually conceptualized as being provided by the market, childcare costs are seen as the key variable in women's decision about participation in the labour market. These models predict lower childcare costs to result in higher female labour force participation (Kimmel, 1998), and several studies provide empirical evidence of this relationship (see Blau and Currie 2006 for a review).

However, many families turn to an informal type of childcare, the childcare provided by relatives, which usually does not involve monetary costs. Even though Heckman called attention to this fact in his pioneering paper on this field (Heckman, 1974), there are very few empirical studies analysing the effect of grandparental childcare in women's work decisions (Aassve et al., 2012; Compton & Pollak, 2014; Dimova & Wolff, 2008, 2011). Grandparental childcare is particularly common in the Mediterranean countries characterized by strong family ties (Dalla Zuanna, 2001). Whereas the percentage of grandparents providing some help with childcare is high throughout all Western Europe, the proportion providing regular childcare (almost weekly or more often) in Southern Europe is almost double that of the Nordic countries (Hank & Buber, 2009). This difference accentuates if one looks at the percentage of grandparents who provide care on a daily basis: around 30% in Italy and Spain *versus* 15% in Germany and Austria and only 2% in countries like Denmark and Sweden.[1] Although we are looking at the Italian case, one might expect grandparental childcare to affect women's work decisions wherever childcare is either not free or not universally available. In fact, the Italian case is particularly interesting because Italy has one of the lowest female labour market participation rates among the OECD countries.

In the present context the topic of this chapter is of particular policy relevance. The European policy goal of delaying retirement to ensure the solvency of pensions plans is accompanied with that of increasing women's labour market participation, set by the Lisbon strategy. However, these two goals are, to some extent, contradictory as delayed retirement might be at odds with the availability of grandparents to provide childcare—especially in countries where a large proportion of the population relies on childcare provided by grandparents (usually grandmothers). In the UK, where around one quarter of parents make use of grandparental childcare (Glaser et al., 2010), there has been pressure for the recognition of the

[1] Authors' calculations from the 2004 wave of the SHARE data.

role of grandparents as providers of childcare. As of 2011, grandparents who give up paid work to provide childcare can claim credits which allow them to qualify for a full basic state pension.

The empirical analysis of the effect of grandparental childcare on female labour force participation is complex due to endogeneity issues. Despite the richness of information concerning work and family, it is not possible to exclude entirely the possibility that there are unobserved preferences related both to childcare and labour market decisions. In these circumstances, if standard OLS regressions were to be used, they would produce biased results. After discussing the most important sources of endogeneity, we use the potential outcome framework (Neyman, 1923; Rubin, 1974) to formalize the problem and to explain why an IV approach should be used instead of standard OLS. As instrumental variables we use information on whether grandparents are alive to exogenously predict the probability of using grandparental childcare. We also discuss what is estimated under different approaches and, in light of this, we contrast our results to the few related empirical works in the literature.

Our results show that the effect of grandparental childcare on mothers' labour supply is positive, statistically significant and economically relevant. Moreover, we find the effect to be heterogeneous: the benefit of grandparental childcare is stronger for mothers with low education, with young children, living in Northern and Central Italy. Differently from other papers using an IV approach, we present tests and robustness checks to support the credibility of our instrumental variables.

2 Data and Variables

The data we use are drawn from the Multiscopo—Famiglie e Soggetti Sociali (Families and Social Subjects) survey,[2] collected in 2003, and released by the Italian National Institute of Statistics (ISTAT). Multiscopo is the main source of information in Italy on the structure and social characteristics of the family, which makes it particularly well suited to study the relationships among family members (including with those living outside the household), childcare decisions regarding young children and leaving home decisions of older children and marriage and fertility intentions, for example. Regarding grandparental childcare, a key advantage of Multiscopo

[2] The non-response rate is about 12%. For further details on the Multiscopo surveys, see ISTAT (2006).

is that it not only provides information on its use but also on its frequency. Therefore, it is possible to know who uses grandparental childcare on a regular basis—which we define as at least twice a week.[3] Unlike other surveys that have been used to study grandparental childcare, Multiscopo includes mothers whose parents have deceased, rendering our identification approach feasible.

Multiscopo collected information on 19,227 Italian families of which 4206 have at least one child younger than 14 years old. Only mothers in a couple who had at least one child younger than 14 years old (3869 observations) were selected. Single mothers were excluded due to the greater complexity in the reconciliation of family and work they face compared to mothers in couples.[4] Ideally we would treat single mothers and mothers in a couple separately, instead of excluding the former, but there are not enough cases. After dropping the very few observations with missing values in each of the variables used in the model, the final sample is composed of 3852 mothers.

The employment status of the mother is defined as working or not. As we can see from Table 1, the labour force participation rate is 66% among mothers that use grandparental childcare and only 42% for the others.

Another important advantage of Multiscopo is that it contains plenty of information on the middle generation (the mother and her partner) and their children, unlike SHARE for example.[5] This allows us to consider a set of control variables concerning the mother, the father and the whole household such as age and age squared of the mother, education of each of the parents (dummy variables indicating tertiary and secondary education; less than secondary education is the reference category) and two variables summarizing the fertility history of the couple: the number of children under 14 years old and whether the youngest one is in pre-school age (3 years old or less).

Availability, quality and cost of formal childcare vary substantially across the Italian regions (Del Boca, 2002). These factors might influence the probability to receive grandparental childcare and can also be associated

[3] We tried different categorizations but the main results (Sect. 6) remain qualitatively the same.

[4] For an overview on single motherhood and childcare choices see Raeymaeckers et al. (2008).

[5] We cannot use SHARE since our instrumental variable is not available in this survey. Moreover, in SHARE we would have information only about two grandparents instead of four, which would provide an incomplete picture of the care of the children in the family.

Table 1 Descriptive statistics

Variables	*All mothers*		*Grandparental childcare: yes*		*Grandparental childcare: no*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Dependent variable						
Mother is working	0.52		0.66		0.42	
Independent variables						
Mother's age	37.20	6.10	36.0	5.50	37.90	6.40
Mother's education: tertiary	0.11		0.13		0.10	
Mother's education: secondary	0.45		0.51		0.41	
Father's education: tertiary	0.11		0.11		0.11	
Father's education: secondary	0.39		0.43		0.37	
At least 1 child aged 0–2	0.29		0.34		0.25	
Number of children younger than 14	1.50	0.63	1.55	0.63	1.46	0.63
Endogenous variable						
Grandparental childcare	0.40		–		–	
Instruments						
Maternal grandmother alive	0.88		0.94		0.83	
Maternal grandfather alive	0.70		0.75		0.66	
Paternal grandmother alive	0.83		0.89		0.78	
Paternal grandfather alive	0.61		0.69		0.56	
Observations	3852		1523		2329	

with labour market conditions (e.g. unemployment rate), which in turn could influence female labour force participation. Moreover, Jappens and Van Bavel (2012) find that mothers' child care choices are related to the family practices and norms prevailing in their region of residence and these vary within the same country. Therefore, we also include in the analyses a set of dummy variables for the region of residence.

Table 1 presents descriptive statistics for the variables used in the analysis both for the whole sample and, separately, for mothers using and not using grandparental childcare. The sample is composed of relatively young mothers (average age is 37 years old) whose parents are, by and large, still alive. There are no noticeable differences in the average characteristics of the two sub-samples (mothers helped and mothers not helped by grandparents), except for the percentage of mothers working. Participation in the labour market is higher among mothers helped in child care by grandparents. Mothers who are not helped are somewhat older and have older children, but the differences are small.

3 The Endogeneity of Grandparental Childcare

Motherhood certainly affects the value women place on their time at home (Blau & Ferber, 1992) but the extent to which this happens depends on women's preference regarding the "externalness" of childcare (El-Attar, 2007), that is, the stronger is the preference for taking care of the children themselves (the most internal type of childcare) the higher is the reservation wage, that is, the minimum wage she would be willing to accept to participate in the labour market. This preference type is probably most often observed in contexts where family ties are strong and where wages and employment rates are low (Alesina et al., 2010). On the other hand, women who have a stronger preference towards participating in the labour market must be willing to use a more external type of childcare. Within this group of mothers there might still be substantial heterogeneity. If in some cases, grandparental childcare is considered as a substitute for formal childcare—a cheaper one—there might be other reasons leading parents to have grandparents looking after their children. For example, concerned about childcare quality, some parents might have a preference against it (Parera-Nicolau & Mumford, 2005; van Ham & Büchel, 2004) and would not use non-relative care even if it was free (Ermisch, 2002). Hence, while some working mothers might choose the next more internal childcare type, grandparental childcare, others might trust formal childcare and send their children to this most external type of childcare—preferring it to grandparental childcare on the grounds of the benefits of formal childcare with respect to child development, for example (Brilli et al., 2011; Felfe & Lalive, 2010; Havnes & Mogstad, 2011).

Figure 1 summarizes graphically the previous discussion. Let us start by assuming that the preference for external childcare is positively associated with mother's willingness to participate in the labour market. Then, consider two groups of mothers who both have preferences "against" grandparental childcare: group A, which includes those with a strong preference for caring for the children themselves, and group B that includes those who prefer formal childcare. In between these two extremes are mothers who are willing to accept help in childcare from grandparents. Up to a certain threshold, τ, the preference for external types of childcare is positively associated with the probability to accept grandparental childcare because mothers accept to externalise childcare to grandparents.

After this point, the relationship turns negative because for high levels of "externalness" mothers start preferring formal to informal childcare. As

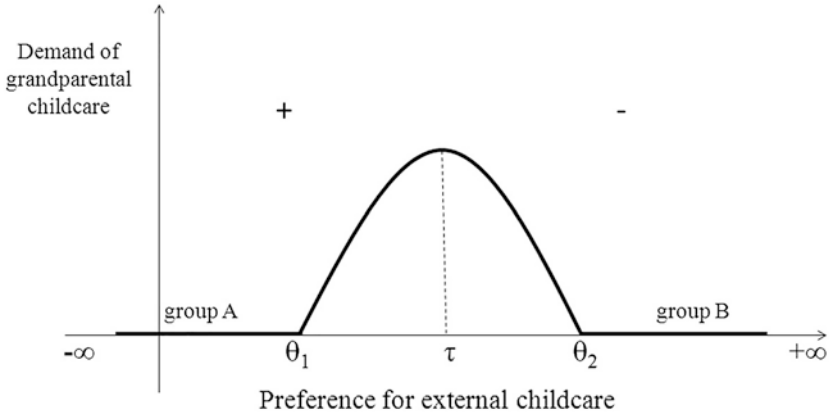

Fig. 1 The relationship between mothers' external preference for childcare and demand of grandparental childcare. (Note: Preference for external childcare represents a latent variable that drives childcare choices. We observe different values of the variable "demand of grandparental childcare" according to the values of this latent variable. For example, if the preference for external childcare is lower that θ_1 or higher than θ_2, then the demand for grandparental childcare will be zero)

a result, unobserved preferences introduce a positive bias in the relationship between observed grandparental childcare and labour market participation for values of "externalness" below τ. The sign of the bias turns negative for values above τ.

The combined effect of the positive correlation between "externalness" and labour supply and the non-linear effect on grandparental childcare will produce an overall upward or downward bias depending on which effect prevails: that of women with a preference for the most internal type of childcare or that of women who prefer an external type of childcare.

Another potential source of bias are grandparents' unobserved characteristics. Grandparental childcare depends on their willingness and ability to do so. It might be that in some cases no childcare is provided by grandparents because they are not willing to forgo their leisure time or retire early in order to take care of their grandchildren. These less traditional grandparents are likely to have less traditional children as well and, consequently, their female children are likely to be participating in the labour market. Indeed, Farré and Vella (2013) find that women's attitudes significantly affect her children's views towards working women and that this

cultural transmission influences female labour market decisions. If these unobserved preferences are not taken into account, the effect of grandparental childcare in mother's labour market participation might be downward biased.

4 Methods

We use the potential outcome framework (Neyman, 1923; Rubin, 1974) to formalize the statistical issues involved in estimating the effect of grandparental childcare on the labour force participation of mothers. We consider a binary treatment, T (= 1 for women who use grandparental childcare on a regular basis, = 0 otherwise), and a binary outcome, Y (= 1 for working mothers, = 0 otherwise). Each mother has two potential outcomes: Y_1, her working status would she have used grandparental childcare (treated) and Y_0, her working status would she have not used it (control). If treated and control mothers would only differ with respect to observed variables, a simple regression model could be used to estimate the effect of interest:

$$Y = \alpha + \delta T + \sum_k \beta_k X_k + \varepsilon, \tag{1}$$

where $Y = Y_1 T + Y_0 (1-T)$ is the observed outcome, α, δ and β represent parameters to be estimated (δ being the one of interest) and ε represents the error term. However, as discussed in the previous section, unobserved preferences for grandparental childcare and labour force participation might be correlated. In this case, the error term in Eq. (1) would be correlated with T, implying the OLS estimator to be biased. We address this problem by instrumenting grandparental childcare.

The instruments we use are whether the mother's parents and her partner's parents are still alive or not. Consequently, we have four instruments. For illustration purposes, let us first consider only one grandparent per woman, that is, only one instrumental variable, Z (= 1 if the grandparent is alive, = 0 otherwise). An instrumental variable (IV) has to affect the probability of using grandparental childcare (relevance) without having a direct influence on the mother working status (validity or exclusion restriction). The exclusion restriction means that any effect of the instrument on the outcome must operate via its effect on the treatment: $Y_{tz} = Y_t, \forall\, t, z \in \{0,1\}$ (Angrist et al., 1996).

An IV can be thought of as a randomization device: the instrument, possibly conditional on control variable X, assigns units to either the treatment or control groups independently of the outcome. Formally,

$$\left(Y_1,\ Y_0,\ T_1,\ T_0\right) \perp Z \mid X, \tag{2}$$

where T_1, T_0 are the potential treatment statuses depending on the instrument's level assignment. In our case, receiving the treatment results from a joint decision of parents and grandparents. Being alive is a necessary but not sufficient condition for grandparents to provide childcare. Grandparents in poor health might have been willing to provide childcare to their grandchildren if only they were able to do so. In what follows, we will first assume that when grandparents are alive they are in sufficiently good health and willing to provide childcare and discuss the bias due to unobserved mothers' preferences for childcare.

To better understand what different estimators estimate, it is useful to employ Angrist et al.'s (1996) classification of units accordingly to the way they may react to the assignment to the different values of the instrument. Panel A of Table 2 shows that a first group, that Angrist et al. (1996) defined as *compliers*, consists of units that are induced to take the treatment by assignment to the treatment, that is, they conform to the instrument assignment ($T_1 = 1$ and $T_0 = 0$). Other units are not influenced by the instrument assignment and are defined either as *always-takers* ($T_1 = T_0 = 1$) if they always take the treatment irrespective of the

Table 2 Classification of mothers by instrument assignment and treatment intake

Grandparent alive (Z)	*Panel A*		*Panel B*	
	Grandparental childcare (T)		Grandparental childcare (T)	
	Yes (=1)	No (=0)	Yes (=1)	No (=0)
Dead (=0)	Defiers/always takers	Compliers/never takers		Compliers/never takers
Alive (=1)	Compliers/always takers	Defiers/never takers	Compliers	Never takers

For simplicity, in this table we consider only one grandparent. Panel A refers to the classification of units (mothers) using the general classification by Angrist et al. (1996). Panel B indicates how this classification simplifies in our context (neither defiers nor always takers exist)

assignment, or *never takers* ($T_1 = T_0 = 0$) if they never take the treatment. Finally, *defiers* do the opposite of their assignment status ($T_1 = 0$ and $T_0 = 1$).

In our case, compliers are mothers who use grandparental childcare when the grandparent is alive and, of course, do not use it when the grandparent is dead. Always-takers and defiers clearly do not exist (T_0 is necessarily equal to 0) and the general classification simplifies into the one described in Panel B of Table 2.[6]

We can think of the endogeneity problem in Eq. (1) as deriving from the fact that mothers' childcare preferences may break the randomization assignment process related to the presence/absence of grandparents. This would be the case of mothers who have preferences such that they would never use grandparental childcare (never takers). OLS is biased because the treated group (composed only by compliers) is compared with a control group which is a mixture of compliers and never takers—see Table 2 panel B. Never takers correspond to the women who have preferences towards each of the extremes of the externalness scale, that is, a preference for the most internal type of childcare (group A in Fig. 1) and a preference for the most external type of childcare (group B in Fig. 1). These two types of never takers give rise to biases in opposite directions, so it is not possible to say a-priori whether the coefficient estimated by OLS is down or upward biased.

So far it has been assumed that when the grandparent is alive he/she is willing to provide childcare. However, in our application the treatment intake depends also on the grandparent's preferences for providing childcare. Less traditional grandparents might prefer to enjoy their free time in social activities rather than provide childcare, for example. In other words, the treatment intake results from a joint decision: mothers must be willing to use grandparental childcare and the grandparent must be willing to provide it. As a consequence, never takers include those mothers who would refuse grandparental childcare if available (grandparent alive and willing to help) and those who would accept help from the grandparent who, however, is not willing to help. The preferences towards childcare of

[6] This situation is similar to the Zelen's (1979) single-consent design where patients are randomly assigned to two groups: the first group of patients is only offered with the standard therapy, while patients in the second group can choose between the standard and new therapy.

this latter group are more similar to those of the compliers than to those of the other never takers. If we assume for a moment that mothers are always willing to use grandparental childcare if available, then never takers would only be composed by those with grandparents not willing to provide childcare. If this were the case, mothers who are compliers or never takers would not differ according to their childcare preferences but to that of their parents. Even in this case, there might be an endogeneity issue as less traditional grandparents are also more likely to have less traditional daughters who might be more likely to be participating in the labour market. In these circumstances, OLS would be biased downward because never takers will be more likely to have strong preferences towards work.

Let us now consider what an IV analysis estimates in our case. Under the exclusion restriction and relevance assumptions, Angrist et al. (1996) show that the IV estimand can be seen as the causal effect on the subpopulation of compliers (local average treatment effect, LATE):

$$\frac{E\left[Y_i \middle| Z_i = 1\right] - E\left[Y_i \middle| Z_i = 0\right]}{E\left[T_i \middle| Z_i = 1\right] - E\left[T_i \middle| Z_i = 0\right]} = E\left[Y_{1i} - Y_{0i} \middle| T_{1i} - T_{0i} = 1\right] = LATE. \quad (3)$$

Angrist et al.'s (1996) result is also based on the monotonicity assumption, which amounts to assume the non-existence of defiers. As noticed before, in our context this assumption is satisfied by the nature of our instrumental variables.[7] Similarly, in our context always-takers do not exist. Consequently, all treated are compliers and the LATE coincides with the average causal effect on the treated (ATT) (Angrist, 2004):

$$\text{LATE} = E\left[Y_{1\text{i}} - Y_{0\text{i}} \middle| T_{1\text{i}} - T_{0\text{i}} = 1\right)\Big] = E\left[Y_{1\text{i}} - Y_{0\text{i}} \middle| T_{\text{i}} = 1\right)\Big] = \text{ATT}$$

An alternative approach to regression model (1) would be to use Z, instead of T, as main independent variable of interest. This approach, known as intention-to-treat (ITT), is justified under assumption (2) and

[7] Angrist et al. (1996) also imposed other two assumptions: (i) the instrument is as good as randomised. This assumption implies that assumption (2) in the text should hold also without conditioning on covariates; (ii) Stable Unit Treatment Value Assumption (SUTVA). SUTVA rules out interference among units: potential outcomes for each unit are not influenced by the level of instrument and treatment assigned to other units.

gives unbiased estimates of the effect of the assignment rather than that of the treatment. Interestingly, the IV estimand will coincide with the ITT estimand divided by the "lease up rate" (the proportion of complying mothers; see Angrist & Pischke, 2009, p. 164):

$$\frac{\text{ITT}}{\pi\left(\text{compliers}\right)} = \frac{E\left[Y_{\text{i}}\middle|Z_i = 1\right)\Big] - E\left[Y_{\text{i}}\middle|Z_i = 0\right)\Big]}{P\left[T_{1\text{i}} - T_{0\text{i}} = 1\right]}.$$

Therefore, we can expect the effect estimated by IV to be higher in magnitude than the one obtained by an ITT-type analysis. The latter underestimates the effect of grandparental childcare because only mothers with grandparents alive can use it and not all mothers with grandparents alive actually use grandparental childcare.

If covariates need to be accommodated to make the assumptions more plausible, as in our case, we can think of Angrist et al.'s (1996) results as applying within cells defined by covariates values (Frölich, 2007). Our situation is more complex also because we have four and not just one binary IV. In the presence of covariates and with many IVs, the IV estimand can be seen as a weighted average of causal effects for instrument-specific compliers (Angrist & Pischke, 2009) but the main insights from the previous discussion still apply. We use a standard two-stage least squares (2SLS) estimator despite the fact that both the endogenous and the dependent variables are binary. This choice is justified by the fact that the interpretation of results is more straightforward (Hellevik, 2009; Powers & Xie, 2008), and tests on the IVs can be easily implemented.[8] In addition to the *F*-test of relevance, we implement a test of over-identification restriction. This test uses the availability of more instruments than endogenous regressors to test if excluded instruments are correctly excluded from the outcome equation (validity of instruments). The tests will be presented in Sect. 6 together with the results.

[8] The use of standard 2SLS methods with dichotomous dependent variables is very common in empirical works. See, for example, Conley and McCabe (2012). We also used a bivariate probit model with the same exclusion restrictions as used in the analyses reported in the paper. Results are very similar to those presented here. We provide further details on this robustness check in Sect. 6.2.

5 Existing Studies on the Effect of Grandparental Childcare on Mothers' Labour Supply

There is limited evidence on the relationship between the use of grandparental childcare and mothers' labour supply. Due to data constraints, some papers use proxy measures for grandparental childcare. In her paper on the effect of (formal and informal) childcare availability on work and fertility decisions in Italy, Del Boca (2002) goes round the endogeneity problem by using as a key variable 'having at least one grandparent alive'—which is assumed to be correlated with the use of grandparental childcare but not to affect directly women's work and fertility decisions. Hers is an ITT analysis which, for the reasons discussed above, is expected to underestimate the effect of grandparental childcare. Nonetheless, Del Boca finds a positive effect of having at least one grandparent alive on the probability that women work.

Ogawa and Ermisch (1996) find a positive relationship between intergenerational co-residence and the labour supply of young women and interpret this result as reflecting the childcare role played by the woman's parents or parents-in-law. Ogawa and Ermisch's analysis is somehow intermediate between an ITT analysis using indicators for alive grandparents and our analysis using actual help received by grandparents. Differently from Ogawa and Ermisch (1996), Abendroth et al. (2012) using data from the second round of the European Social Survey (ESS) find no significant effect of having at least one of the grandparents living in the household on the number of working hours of employed mothers.

García-Morán and Kuehn (2012) find that having grandparents close increases the probability of mothers to work. However, the German Socio-Economic Panel (GSOEP) employed by the authors does not provide information on grandparental childcare but only on childcare provided in general by relatives. These authors adopt a simulation approach using a general equilibrium model of residence choice, fertility decisions and female labour force participation that accounts for the interrelationships among these processes. They simulate alternative counterfactual scenarios and find that without grandparent-provided childcare there would be fewer women participating in the labour market.

Compton and Pollak (2014) show that close geographical proximity to mothers or mothers-in-law has a substantial positive effect on the labour force participation of women with young children. They argue that the mechanism through which proximity increases labour supply is the

availability of childcare. The authors argue that proximity is a good instrumental variable for childcare arrangements as the positive effect of proximity on labour force participation does not extend to groups for which grandparent-provided childcare is not a determinant for their labour supply like men, unmarried women without children or those with sick mothers or sick mothers-in-law.

The distance between mothers and grandparents' residence strongly influences the probability to receive childcare (see e.g. García-Morán & Kuehn, 2012). However, proximity of mothers and their parents is related to family values (Hank, 2007) which, in turn, might be associated with unobserved preferences about female labour force participation. The falsification tests implemented by Compton and Pollak (2014) seem to show that there is no clear evidence that family values are associated with labour force participation preferences. However, when forming a stable union, mothers may choose their geographical location anticipating the need or desire of using grandparental childcare. Along these lines, Leopold et al. (2012) find that young adults who had an own child tend to stay closer to their parental home. The falsification tests implemented by Compton and Pollak (2014) cannot rule out this issue. Thus, we claim that the geographical distance might not be a credible instrument. Nonetheless, in Sect. 6.3 we will use information on geographical distance between the mother and grandparents in an alternative definition of our instrumental variables.

Data on grandparental childcare from ten European countries is used by Dimova and Wolff (2011) who deal with the endogeneity of grandparental childcare in two ways. First, exploiting data available on several daughters of the same grandmother they use a fixed-effects analysis to control for unobserved household-level characteristics. As mothers' specific unobserved confounders might still have an impact, the authors also jointly estimate three equations for mothers' working decision and the probabilities to receive time and money transfers from grandmothers—allowing the error terms in the three models to be correlated. The authors exclude some variables related to grandparents' characteristics from the mothers' labour supply equation using them implicitly as instrumental variables. For example, they exclude grandparents' age and health status that might well be factors influencing the daughter's decision to work or not.

The effect estimated by the Dimova and Wolff (2011) is an average effect from people residing in the ten countries included in the study. As

noticed by Aassve et al. (2012), this approach might be problematic insofar as the estimated average impact might represent poorly the country-specific effects—with important consequences for the interpretation of the results in terms of policy implications. Similarly to Dimova and Wolff (2011), Aassve et al. (2012) use a bivariate probit model for mothers' labour force participation and grandparental childcare using information on whether the respondent's mother is alive and the number of siblings the respondent has as instrumental variables. The first instrument they use is similar to ours, but they limit their focus to the respondent's mother because the GGS does not provide information on the parents of the respondent's partner. Their second instrument could be problematic if the size of the respondent's family is correlated to unobserved preferences for childcare and work.

The main result in Dimova and Wolff (2011) is that downward time transfers have a strong and positive effect on mother's labour force participation. A similar result is found by the same authors (Dimova & Wolff, 2008) on the particular population of immigrants in France and by Aassve et al. (2012) but only for some of the GGS countries.

6 Results

OLS, ITT and 2SLS Estimates of the Effect of Grandparental Childcare on Mothers' Labour Force Participation

We start by presenting the OLS and ITT results so that later we can compare them with the results of the IV approach. Table 3 presents the results of the estimation of a simple linear probability model (OLS) where mother's participation in the labour market is regressed on grandparental childcare and on the control variables described in Sect. 2. Since our IV approach uses as instruments having or not grandparents alive, we cannot include information on the grandparents. Characteristics of grandparents such as the health status and geographical proximity are only used in some of the robustness checks presented in the next section. Although the characteristics of grandparents would be essential to study the determinants of grandparental childcare (see e.g. Hank & Buber, 2009), this is beyond the scope of this chapter.

We find a highly significant positive association between the use of grandparental childcare and the probability to work. Mothers who use grandparental childcare are 20 percentage points more likely to work. The

Table 3 OLS estimates of the effect of grandparental childcare on mother's work

Dependent variable: mother is working	*Coef.*	*Robust SE*	*Sig.*
Mother's age	0.045	0.011	***
Mother's age squared	–0.001	0.000	***
Mother's education: tertiary	0.409	0.027	***
Mother's education: secondary	0.221	0.017	***
Father's education: tertiary	0.022	0.028	
Father's education: secondary	0.008	0.017	
At least 1 child aged 0–2	0.007	0.018	
Number of children younger than 14	–0.061	0.011	***
Grandparental childcare	0.198	0.015	***
Constant	–0.481	0.200	**
Observations	3852		

Robust standard errors; *, ** and *** significant at 10, 5, and 1% levels, respectively. Dummy variables for regions are included but coefficients are not reported

effects of other characteristics of the mother are as expected: the number of children and lower education decrease the probability of working while father's education turns out to be not significant.

Table 4 shows the results of a linear probability model where the covariates of interest are the variables that will be used as instruments in the IV approach. From this type of ITT analysis we can see that having grandparents alive has a positive effect on the probability of mothers to work.

The results of the 2SLS estimation[9] in Table 5 show that mothers who use grandparental childcare are 32 percentage points more likely to work, an effect comparable to that of having tertiary versus primary education. To better appreciate the magnitude of the estimated effect, we calculated the predicted probability of working of a typical woman (aged 35, with secondary education, with one child aged 0–2) who uses grandparental childcare and that of a similar woman who does not use grandparental childcare. These are 81.6 and 49.4, respectively.

The estimated effect of grandparental childcare is significant and higher than the one estimated by OLS. As argued in Sect. 4, this OLS downward bias might be due to the prevailing effect of mothers with a preference for the most external type of childcare with respect to the effect of mothers with a preference for the most internal type of childcare and/or to the effect of grandparents' unobserved preferences.

[9] We used the ivreg2 command in STATA (Baum et al., 2007).

Table 4 OLS estimates of the effect of grandparents being alive on mother's work

Dependent variable: mother is working	*Coef.*	*Robust SE*	*Sig.*
Mother's age	0.053	0.011	***
Mother's age squared	–0.001	0.000	***
Mother's education: tertiary	0.432	0.027	***
Mother's education: secondary	0.237	0.018	***
Father's education: tertiary	0.013	0.029	
Father's education: secondary	0.006	0.018	
At least 1 child aged 0–2	0.009	0.018	
Number of children younger than 14	–0.056	0.012	***
Maternal grandmother alive	0.037	0.024	
Maternal grandfather alive	–0.008	0.017	
Paternal grandmother alive	0.043	0.021	**
Paternal grandfather alive	0.027	0.016	*
Constant	–0.615	0.204	
Observations	3852		

Robust standard errors; *, ** and *** significant at 10, 5, and 1% levels, respectively. Dummy variables for regions included but coefficients not reported

Table 5 Two-stage least squares (2SLS) estimates of the effect of grandparental childcare on mother's work

Dependent variable: mother is working	*Coef.*	*Robust SE*	*Sig.*
Mother's age	0.040	0.012	***
Mother's age squared	0.000	0.000	**
Mother's education: tertiary	0.392	0.031	***
Mother's education: secondary	0.208	0.020	***
Father's education: tertiary	0.026	0.028	
Father's education: secondary	0.006	0.017	
At least 1 child aged 0–2	0.005	0.018	
Number of children younger than 14	–0.065	0.012	***
Grandparental childcare	0.323	0.106	***
Constant	–0.448	0.203	**
Tests on the instrumental variables			
Cragg–Donald Wald *F* statistic	21.506		
Hansen *J* statistic	2.244		
($\chi^2(3)$ *p* value)	(0.523)		
Observations	3852		

Instrumental variables 2SLS linear regression (robust standard errors: *, ** and *** significant at 10, 5, and 1% levels, respectively). Instrumented variable: grandparental childcare. Instrumental variables: maternal grandmother alive, maternal grandfather alive, paternal grandmother alive, paternal grandfather alive. Dummy variables for regions included but coefficients not reported

As expected, the effect estimated by 2SLS is also stronger than the one obtained with the ITT analysis (Table 4). Our estimates suggest that the effect of grandparental childcare obtained using proxy variables for provision of grandparental childcare, as in Del Boca (2002), might be underestimated. Similarly, García-Morán and Kuehn (2012) find that the effect of proximity on labour force participation is smaller than the effect of actual childcare provided by relatives.

IV Assumptions: Tests and Robustness Checks

As discussed in Sect. 4, the IV approach is based on some fundamental assumptions of which monotonicity, IVs relevance and validity are the most important. As already noticed, monotonicity is satisfied by the nature of the instrument.

Descriptive evidence in favour of the IVs relevance, that is, non-zero association between the instrumental variables and the endogenous variable, is provided in Table 1 which shows that the proportion of grandparents alive is considerably higher among mothers receiving help. The relevance of the instruments is confirmed by the results in Table 6. The

Table 6 OLS estimates of the effect of grandparents' being alive on grandparental childcare

Dependent variable: grandparental childcare	*Coef.*	*Robust SE*	*Sig.*
Mother's age	0.040	0.011	***
Mother's age squared	–0.001	0.000	***
Mother's education: tertiary	0.124	0.030	***
Mother's education: secondary	0.090	0.018	***
Father's education: tertiary	–0.040	0.030	
Father's education: secondary	0.004	0.018	
At least 1 child aged 0–2	0.014	0.020	
Number of children younger than 14	0.030	0.013	**
Maternal grandmother alive	0.154	0.021	***
Maternal grandfather alive	0.022	0.018	
Paternal grandmother alive	0.084	0.020	***
Paternal grandfather alive	0.057	0.017	***
Constant	–0.513	0.216	**
Observations	3852		

Robust standard errors; *, ** and *** significant at 10, 5, and 1% levels, respectively. Dummy variables for regions included but coefficients not reported

four instrumental variables, in particular the ones regarding grandmothers, are significantly associated with grandparental childcare after having controlled for the effect of covariates.

The fact that the instruments are not weak is also confirmed by the Cragg-Donald Wald F-test[10] presented in Table 5. The F-statistic of the test of correlation among the instruments and the endogenous variable is 21.506 overcoming the threshold of 10 usually considered acceptable (Staiger & Stock, 1997).

We turn now to the discussion of the exclusion restriction. This assumption would be violated if the instrumental variables influence the outcome through channels other than childcare provision. The death of a grandparent might produce an income shock, for example. If a mother relies on grandparental financial support to be out of the labour market, their death might push her to find a job. On the other hand, working mothers that receive an inheritance after grandparents' death might decide to quit working. Also the health status of grandparents might be associated with mothers' probability to work. For example having grandparents alive but in poor health might have a negative impact on the probability of mothers to work if they have to take care of the elderly. Having more IVs than endogenous variables allows us to test the exclusion restriction assumption by implementing a test of over-identifying restrictions. The joint null hypothesis that the instruments are valid and that the excluded instruments are correctly excluded from the estimated equation is not rejected (Hansen J-statistic = 2.244; p-value = 0.523). We also implemented a simple falsification test. We estimated the effect of having grandparents alive on the working decisions of mothers without children. We selected a sample of 2919 women in a couple but without children, in the same age-range as mothers in our main sample (between 23 and 51 years old). We find that none of the four IVs was significantly associated with Υ (results available from the authors upon request). This provides evidence that having or not grandparents alive does not have an effect per se on the probability of mothers to work but only through its effect on the provision of grandparental childcare.

Table 7 shows the results of different robustness checks. Our IV approach would not be valid if having grandparents alive has a direct effect on mothers' labour market outcomes. That could be the case if

[10] This variant of the standard Cragg-Donal F-test is required because we are allowing *heteroskedasticity* of any kind and so errors to be not i.i.d. (Baum et al., 2007).

Table 7 Robustness checks

Mother's characteristics	*OLS*			*2SLS*			*Obs.*
	Coef.	*Robust SE*	*Sig.*	*Coef.*	*Robust SE*	*Sig.*	
Analysis 1							
Only families with all grandparents in good health and without need of assistance	0.184	0.018	***	0.238	0.115	**	2851
Analysis 2							
Instrumental variable equal to 1 if the grandparent is alive and geographically close	0.198	0.015	***	0.304	0.052	***	3852
Analysis 3							
Criteria in both analysis 1 and analysis 2	0.184	0.018	***	0.309	0.058	***	2851

*, ** and *** significant at 10, 5, and 1% levels, respectively. Outcome: mother's working status; Instrumented variable: grandparental childcare; Instrumental variables: maternal grandmother alive, maternal grandfather alive, paternal grandmother alive, paternal grandfather alive. All other control variables, as shown in Table 5, are included but coefficients are not reported

grandparents who are alive but in bad health need to be taken care of. So, in a first robustness check we exclude all families where at least one grandparent is chronically sick, assisted or helped by the mother.

In a second robustness check we change the definition of the instrumental variables. Each instrumental variable is now equal to 1 if the correspondent grandparent is alive *and* geographically close (i.e. if the distance is at most 16 km). Finally, we selected only grandparents in good health conditions *and* used the new instrumental variables. The estimated effect of grandparenting remains similar to the ones obtained in the analyses shown before.[11]

Finally, we checked if using a probit model we get similar results than those presented here.[12] To account for endogeneity we then considered a bivariate probit model where we jointly estimated two probit equations. In the first, the outcome was mother's working status and it was regressed on grandparental childcare and all the control variables. In the second, the

[11] We tried additional specifications of the IVs obtaining similar results. In particular, we considered the number of grandparents alive and we restricted the analyses only to grandmothers. Results are available from the authors upon request.

[12] Results not shown (available upon request).

probability of receiving grandparental childcare was estimated as function of our instrumental variables and all control variables. To account for endogeneity, the errors in the two equations were correlated and we used as exclusion restrictions information on grandparents being alive or not (similarly as done in the 2SLS). The estimated coefficient of grandparental childcare in the mothers' labour force participation equation is positive and statistically significant both in the probit and bivariate probit analysis and is higher in the second case. So, this confirms the downward bias we obtain when endogeneity is not taken into account. Finally, we calculated the size of the effect of grandparental childcare on women's labour force participation on the probability scale obtaining a result (30%) very similar to the estimated effect using 2SLS (32%).

Heterogeneity of the Grandparental Childcare Effect

In this section we address the possibility that grandparental childcare has a different effect for different subgroups of mothers. Exploiting the large sample size, we divide it, in turn, according to the level of education of the mother, the number and age of the children in the household and geographical area of residence. Table 8 shows the results both using the OLS and the IV approach. Again, IV estimates are larger than the OLS ones with the exception of mothers with tertiary education for whom the effect is not significant. However, for this group the sample size is rather small. For all the other sub-groups the IV estimate is significant. The effect of grandparental childcare is stronger for those mothers who need childcare the most: mothers with lower education (that can signal lower income, and so lower availability of resources for alternative types of childcare) and mothers with younger children.

The pattern of differences between OLS and 2SLS across the educational groups can be due to less heterogeneity within the group of more educated women in what regards "externalness" preference. In the extreme case in which all more educated women share the same preferences, the bias due to not observing them would not exist. The same does not seem to happen for lower educated women. The OLS results for this group are downward biased suggesting that there are important differences among lower educated women in what concerns "externalness" of childcare preferences. A similar reasoning applies with respect to the age of children as most children older than 3 go to formal childcare and so there is less heterogeneity within this group.

Table 8 Heterogeneous effects of grandparental childcare on mother's work. OLS versus two-stage least squares (2SLS) estimates

Mother's characteristics	*OLS*			*2SLS*			*Obs.*
	Coef.	*Robust SE*	*Sig.*	*Coef.*	*Robust SE*	*Sig.*	
Education							
Primary	0.194	0.024	***	0.370	0.141	***	1700
Secondary	0.231	0.022	***	0.284	0.166	*	1724
Tertiary	0.087	0.037	**	0.047	0.267		428
Age of children							
At least one child aged 0–2	0.209	0.027	***	0.563	0.177	***	1103
All children older than 2	0.191	0.018	***	0.192	0.124		2749
Number of children							
One child	0.214	0.020	***	0.261	0.160		2189
More than one child	0.178	0.023	***	0.277	0.141	*	1663
Residence							
North Italy	0.202	0.024	***	0.484	0.159	***	1507
Centre Italy	0.212	0.037	***	0.551	0.235	**	605
South Italy	0.180	0.023	***	0.205	0.174		1740
Working time arrangement							
Full-time	0.196	0.017	***	0.327	0.111	***	3214
Part-time	0.168	0.018	***	0.328	0.136	**	2497

*, ** and *** significant at 10, 5, and 1% levels, respectively. Outcome: mother's working status; Instrumented variable: grandparental childcare; Instrumental variables: maternal grandmother alive, maternal grandfather alive, paternal grandmother alive, paternal grandfather alive. All other control variables, as shown in Table 5, are included but coefficients are not reported

Similarly, comparing the OLS and 2SLS by geographical area of residence, we notice large differences only in the North and Centre, where the 2SLS are much bigger than the OLS ones and they are statistically significant. Similarly to what discussed before, this can indicate less heterogeneity in unobserved preferences within the Southern women but it may also indicate that women's preference in the South may be less important because they are constrained by unfavourable labour market conditions: even though a woman has a strong preference for working and for external types of childcare, she might find herself doing childcare because she cannot find an employment.

We conclude by assessing if the availability of grandparental childcare has a different impact for mothers working full or part time. In the last two lines of Table 8, we report the estimated effect of grandparental childcare care on two samples that exclude, respectively, mothers working part-time

and mothers working full-time. Independently on the number of work hours, grandparental childcare has a strong and similar effect on the probability of working. The patterns of bias suggest that unobserved preferences differ more according to labour market participation than to labour market attachment. Grandparental childcare impacts the probability to work rather than the number of hours worked.

7 Summary and Concluding Remarks

This chapter studies the role of informal childcare provided by grandparents on the reconciliation of motherhood and labour force participation of women. The association between help with childcare provided by grandparents and mother's work decisions is not necessarily causal. Unobserved preferences for a more external or internal type of childcare might be associated with the propensity of mothers to work. As a consequence, the results from a simple regression might be biased. In the framework of the potential outcome approach, this problem can be seen as arising from the non-compliance of mothers to the availability of grandparents. These never-taker mothers are likely to have unobserved preferences different from those of the compliers. Some of them might refuse grandparental childcare because of a strong preference for an external (formal) type of childcare, others because of a preference for the most internal type of childcare (they want to look after the children themselves). The presence of these two groups produces biases in opposite directions and it is not clear a-priori if a naive OLS analysis would be upward or downward biased. Also grandparents' unobserved preferences for providing childcare might be correlated with mothers' propensity to participate in the labour market.

Using an instrumental variable approach on Italian data, our results show that having grandparents helping with childcare increases mother's labour force participation. The effect is stronger for lower educated mothers, with young children and living in the North and Centre.

We argue that the effect of grandparental childcare found in previous studies might be underestimated. Due to data limitations, some studies used proxy variables for grandparental childcare, such as having grandparents alive (Del Boca, 2002) and co-residence with grandparents (Ogawa & Ermisch, 1996). These approaches are similar to Intention-to-Treat analyses where the effect of the assignment and not that of the treatment is estimated. Since it is not always the case that once the conditions for

grandparental childcare provision are met (grandparents willing and able to provide help), this type of childcare is actually used, these studies are likely to underestimate the effect of informal care provided by grandparents.

Our results seem to support the idea that grandparents are more used as a substitute of formal childcare than as a substitute of parental childcare. The fact that all OLS analyses underestimate the effect of grandparental childcare can be interpreted as a prevailing effect of women with a preference for the most external type of childcare. This suggests that a policy directed towards increasing availability of formal childcare would be effective in increasing female labour force participation, especially considering that grandparents may be increasingly less available for taking care of their grandchildren. Insofar as grandparents who are still active in the labour market might be less available to help in looking after their grandchildren, the evidence presented here suggests a need to frame retirement policies in the larger picture of family and gender policies (Brussig & Knuth, 2007). This is particularly relevant because despite claiming that the provision of high-quality childcare is a priority, the European Commission also acknowledges that most member states have not been able to ensure the desired access to formal childcare (European Commission, 2009). In other words, an increase in retirement age would not be detrimental for female labour force participation only if it is accompanied by a sufficient increase in the availability of childcare. This conclusion is reinforced by the fact that the effect is stronger for low educated women who, earning relatively lower wages, may find it more difficult to pay for the services offered by the private market.

However, more childcare services alone may not be enough to attain the goal of increasing female labour force participation if not accompanied by an increase in its perceived quality. Similarly to what happens when people use many other services, parents face information asymmetry when using childcare services (either privately or publicly provided). This might make mothers not willing to trust their children to institutions even if they have a latent preference for an external type of childcare. Some studies that hint at the importance of trust in the quality of childcare in childcare choices (Borck & Wrohlich, 2008; Hank & Kreyenfeld, 2003; Shlay, 2010 and El-Attar, 2007) find that lower trust decreases the probability of leaving the child with a more external type of childcare. If parents do not trust the quality of formal childcare, an effect of grandparental childcare in mothers' work decisions may be observed even if there were free and universal access to formal childcare.

References

Aassve, A., Arpino, B., & Goisis, A. (2012). Grandparenting and Mothers' Labour Force Participation: A Comparative Analysis Using the Generations and Gender Survey. *Demographic Research, 27*(3), 53–84.

Abendroth, A. K., van der Lippe, T., & Maas, I. (2012). Social Support and the Working Hours of Employed Mothers in Europe: The Relevance of the State, the Workplace, and the Family. *Social Science Research, 41*(3), 581–597.

Alesina, A., Algan, Y., Cahuc, P., & Giuliano, P. (2010). *Family Ties and the Regulation of Labor*. NBER Working Paper 15747. NBER.

Angrist, J. D. (2004). Treatment Effect Heterogeneity in Theory and Practice. *Economic Journal, 114*(494), C52–C83.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association, 91*, 444–472.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2007). Enhanced Routines for Instrumental Variables/GMM Estimation and Testing. *Stata Journal, 7*(4), 465–506.

Blau, D., & Currie, J. (2006). *Preschool, Day Care, and After School Care: Who's Minding the Kids?* In E. Hanushek & F. Welch (Eds.), *Handbook on the Economics of Education, Volume 2, Chapter 20*. North Holland Press.

Blau, F. D., & Ferber, M. (1992). *The Economics of Women, Men, and Work* (2nd ed.). Prentice-Hall.

Borck, R., & Wrohlich, K. (2008). *Preferences for Childcare Policies: Theory and Evidence*. IZA Discussion Paper 3694.

Brilli, Y., Del Boca, D., & Pronzato, C. (2011). *Exploring the Impacts of Public Childcare on Mothers and Children in Italy: Does Rationing Play a Role?* IZA Discussion Paper 5918. IZA.

Brussig, M., & Knuth, M. (2007). Raising Statutory Pension Age: Extending Careers Or Unemployment of the Elderly? Paper Presented at the 5th International Research Conference on Social Security. http://www.issa.int

Compton, J., & Pollak, R. A. (2014). Family Proximity, Childcare, and Women's Labor Force Attachment. *Journal of Urban Economics, 79*(C), 72–90.

Conley, D., & McCabe, B. J. (2012). Bribery Or Just Desserts? Evidence on the Influence of Congressional Reproductive Policy Voting Patterns on PAC Contributions from Exogenous Variation in the Sex Mix of Legislator Offspring. *Social Science Research, 41*, 120–129.

Dalla Zuanna, G. (2001). The Banquet of Aeolus: A Familistic Interpretation of Italy's Lowest Low Fertility. *Demographic Research, 4*, 133–162.

Del Boca, D. (2002). The Effect of Child Care and Part Time Opportunities on Participation and Fertility Decisions in Italy. *Journal of Population Economics, 15*(3), 549–573.

Dimova, R., & Wolff, F.-C. (2008). Grandchild Care Transfers by Ageing Immigrants in France: Intra-household Allocation and Labour Market Implications. *European Journal Population, 24*, 315–340.

Dimova, R., & Wolff, F.-C. (2011). Do Downward Private Transfers Enhance Maternal Labor Supply? Evidence from Around Europe. *Journal of Population Economics, 24*(3), 911–933.

El-Attar, M. (2007). *Trust, Child Care Technology Choice and Female Labor Force Participation*. IZA Discussion Paper 3135. IZA.

Ermisch, J. (2002). Symposium on "Child care". *Journal of Population Economics, 15*, 463–464.

European Commission. (2009). Dealing with the Impact of an Ageing Population in the EU (2009 Ageing Report).

Farré, L., & Vella, F. (2013, April). The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labor Force Participation, *Economica, 80*(318), 219–247.

Felfe, C., & Lalive, R. (2010). *How Does Early Childcare Affect Child Development? Learning from the Children of German Unification, CESifo Area Conference on Economics of Education*. Center for Economics Studies, mimeo.

Frölich, M. (2007). Non Parametric IV Estimation of Local Average Treatment Effects with Covariates. *Journal of Econometrics, 139*, 35–75.

García-Morán, E., & Kuehn, Z. (2012). *With Strings Attached: Grandparent-Provided Child Care, Fertility and Female Labor Market Outcomes*. MPRA Paper No. 37001.

Glaser, K., Ribé Montserrat, E., Waginger, U., Price, D., Stuchbury, R., & Tinker, A. (2010). *Grandparenting in Europe*. Grandparents Plus.

Hank, K. (2007). Proximity and Contacts Between Older Parents and Their Children: A European Comparison. *Journal of Marriage and Family, 69*(1), 157–173.

Hank, K., & Buber, I. (2009). Grandparents Caring for Their Grandchildren: Findings from the 2004 Survey of Health, Ageing, and Retirement in Europe. *Journal of Family Issues, 30*(1), 53–73.

Hank, K., & Kreyenfeld, M. (2003). A Multilevel Analysis of Child Care and Women's Fertility Decisions in Western Germany. *Journal of Marriage and Family, 65*, 584–596.

Havnes, T., & Mogstad, M. (2011). No Child Left Behind. Universal Childcare and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy, 3*, 97–129.

Heckman, J. J. (1974). Effects of Child-Care Programs on Women's Work Effort. *The Journal of Political Economy, 82*(2), S136–S163.

Hellevik, O. (2009). Linear Versus Logistic Regression When the Dependent Variable Is a Dichotomy. *Quality and Quantity, 43*, 59–74.

ISTAT. (2006). Il sistema di indagini sociali multiscopo. Contenuti e metodologia delle indagini, ISTAT Working Paper Series "Metodi e norme" n.31.

Jappens, M., & Van Bavel, J. (2012). Regional Family Norms and Child Care by Grandparents in Europe. *Demographic Research, 27*(4), 85–120.

Kimmel, J. (1998). Child Care Costs as a Barrier to Employment for Single and Married Mothers. *Review of Economics and Statistics, 80*(2), 287–299.

Leopold, T., Geißler, F., & Pink, S. (2012). How Far Do Children Move? Spatial Distances After Leaving the Parental Home. *Social Science Research, 41*(4), 991–1002.

Neyman, J. (1923). On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9. *Translated in Statistical Science, 5*(4), 465–480, (1990).

Ogawa, N., & Ermisch, J. F. (1996). Family Structure, Home Time Demands and the Employment Patterns of Japanese Married Women. *Journal of Labor Economics, 14*, 677–702.

Parera-Nicolau, A., & Mumford, K. (2005). *Labour Supply and Childcare for British Mothers in Two-Parent Families: A Structural Approach*. IZA Discussion Paper 1908. IZA.

Powers, D. A., & Xie, Y. (2008). *Statistical Methods for Categorical Data Analysis* (2nd ed.). Emerald, Howard House, England.

Raeymaeckers, P., Dewilde, C., Snoeckx, L., & Mortelmans, D. (2008). Childcare Strategies of Divorced Mothers in Europe: A Comparative Analysis. *European Sociological Review, 24*(1), 115–131.

Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology, 66*, 688–701.

Shlay, A. B. (2010). African American, White and Hispanic Child Care Preferences: A Factorial Survey Analysis of Welfare Leavers by Race and Ethnicity. *Social Science Research, 39*, 125–141.

Staiger, D., & Stock, J. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65*(3), 557–586.

van Ham, M., & Büchel, F. (2004). *Females' Willingness to Work and the Discouragement Effect of a Poor Local Childcare Provision*. IZA Discussion Paper 1220. IZA.

Zelen, M. (1979). A New Design for Randomized Clinical Trials. *England Journal of Medicine, 300*, 1242–1245.

Youth Labor Market Integration: The Role of Shocks and Institutions

Werner Eichhorst and André Portela

1 Introduction

In general, and globally, there is a stable pattern of higher volatility with regard to employment and unemployment concerning young people related to prime-aged workers, but there is also huge cross-country variation in the extent that young people are neither in education, employment or training both in good times and during crisis periods, as recently highlighted once more during the COVID-19 pandemic (ILO, 2021).

The transition from school to work is heavily influenced by institutional arrangements and economic conditions. If supportive and inclusive, labor market institutions, namely effective vocational training and balanced

W. Eichhorst (✉)
Bremen University, Bremen, Germany
e-mail: Eichhorst@iza.org

A. Portela
Escola de Economia de São Paulo da Fundação Getulio Vargas, Bela Vista, São Paulo, Brazil
e-mail: andre.portela.souza@fgv.br

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_25

employment protection, can shorten the phase of transition and increase the chance of access to a regular job, while a heavily segmented labor market tends to keep young people in longer sequences of vulnerable, often temporary jobs. This difficulty is exacerbated during and after recessions when hiring probabilities decline. This can lead to long-lasting scars in terms of employment and remuneration prospects later in life (see Cockx, 2016 for an overview). Hence, a closer look into the interaction between institutions on the one hand and economic fluctuations on the other hand can shed more light on this relation. However, we distinguish between developed and developing countries given the fundamental differences in the institutional setup and the economic environment.

2 Shocks and Institutions in Youth Labor Markets—European Experiences

The European experience with both the financial crisis and COVID-19 generally shows that youth are less well integrated into the labor market than their adult counterparts, exhibiting higher unemployment and NEET rates during and after crises that take time to recover.

Yet, not all countries are affected equally, pointing at the crucial interaction of economic shocks and labor market institutions. Even under adverse economic conditions, the situation remains quite favorable in some countries, in particular where well-established regular vocational training continue to provide a viable bridge to young people trying to enter the labor market (Eichhorst et al., 2015; Pastore, 2018; Cahuc et al., 2013). In periods of crisis, most of the training contracts tend to be maintained, sometimes encouraged by public subsidies as witnessed in 2020. This is still true despite structural change in the labor market that tends to work against medium-skilled occupations and increasing gaps in professional outlooks between tertiary educated workers and those with (upper) secondary training. Overall, labor market developments in Europe are influenced by the relative weight and prominence of vocational training systems that combine training on the job with real work experience and learning phases at schools. However, it becomes increasingly urgent and obvious to update vocational training systems regularly in order to keep pace with technological and structural change as to keep them attractive for employers and for young workers alike. This might mean a more modular, transferable setup and greater openness towards continued education

at tertiary level. At the same time, modular training would also be more open for more vulnerable young people that may face difficulties in entering regular training paths.

In many countries in Europe, however, economic crises tend to have severe negative effects on the chances of youths to make a transition into regular jobs. In fact, problematic labor market institutions tend to aggravate. For a long time, strict employment protection for open-ended contracts combined with relaxed conditions regarding temporary contracts, agency work and solo self-employment have created a particularly vulnerable secondary labor market segment. Young people could access this segment upon leaving education, but a transition to permanent jobs was difficult even in normal times and became even more problematic and protracted during a recession, that is, implying much longer sequences of temporary contracts and long-term scarring effects. However, the divide between permanent and temporary contracts and fixed-term contracts saw some reforms aimed at closing this regulatory gap over the 2010s (Eichhorst & Marx, 2021), yet, the dualistic structure of labor markets continues to exist and was confirmed again with the reaction to the COVID-19 crisis. Here, to mitigate the impact of crises discretionary measures tend to be more urgent, as shown most recently during the pandemic. Some countries have expanded, on an ad hoc basis, access to unemployment benefits for temporary workers and solo self-employed, and they have started to provide additional incentives to employers to encourage hiring and training (OECD, 2020, 2021).

One of the most prominent initiatives taken after the financial crisis and renewed in 2020 is the EU youth guarantee (Escudero & Mourelo, 2015, Eichhorst & Rinne, 2017), establishing a political obligation to provide a job or training offer to all young people within four months of unemployment or leaving education. Despite the provision of considerable funds, the youth guarantee suffered from implementation bottlenecks in countries with high youth unemployment. Typically, discretionary interventions cannot easily overcome structural conditions that work against a smooth youth integration into work, and structural reforms take time to reverse long-standing dualistic employment models. However, training incentives based on programs with subsidies or incentives to maintain vocational training at firm level despite economic uncertainty can help stabilize existing arrangements.

3 Policies and Institutions—Developing Countries Experiences

The economic crisis due to COVID-19 hit the affected youth cohort twice in developing countries. First, it exacerbated the challenges of generally deficient education systems and, second, it is a negative shock in a situation characterized by already volatile youth labor market transitions. These two impacts can have lasting effects on them as is true in the European framework, but with notable differences.

Early in life, young individuals in developing countries face two possibilities to continue their human capital accumulation process. They can either advance in their formal education or start to engage in labor market activities and gain on the job training (or a combination of both). The choice between work and study depends on many factors such as (i) the availability and access of formal education and its quality; (ii) the availability and access of vocational and professional training and its quality; (iii) the labor market conditions and the availability of on the job training; and (iv) the economic conditions of their families.

Investment in education is expected to increase general human capital accumulation. However, these gains can only be realized if the quality of education is such that the accumulation actually happens and if individuals can finance the investment using their own economic resources or have access to credit. Individuals from vulnerable families with low resources or credit constraints are most likely unable to realize these investments or have access to low-quality education only.

As a consequence, many individuals rather rely on earlier job market experience to overcome these difficulties. Although youth labor can be harmful—particularly in the long run—by hindering the acquisition of formal education (both quantitatively and qualitatively), it may be beneficial to a young laborer by providing vocational training from learning by doing and general workplace experience as well as the potential for making contacts, learning job market strategies, and so on. Furthermore, youth labor could be a way to finance education that an individual would not otherwise have access to. In fact, there are some evidences of positive long-term effect of youth labor (e.g., Emerson and Souza [2011]).

The usual policy options range from policies that impose constraints to the school-work choice set to a variety of policies that expand this set of choices. Among the policies that impose constraints are the minimum employment age laws (MEAL) and the compulsory schooling laws (CSL).

However, the evidence of the MEAL in developing countries is scarce. So far it suggests that its short-run effect on youth labor and schooling depends on the context. This is so due to the weak enforcement of the law and the prevalence of informal labor market and the different types of work such as non-paid work, family farmers, etc. (e.g., Edmonds & Shresta, 2012).

Among policies that increase the choice set are measures aimed to improve the quality of formal education and policies that expand vocational and professional education. The evidence for developed countries suggests that interventions in high school education may produce positive effects and be cost effective. Reasonably, high school interventions may not be too late (e.g., Angrist et al. [2006]). Moreover, there has been an expansion of technical and vocational education and training (TVET) programs. Not only may these programs build occupational-specific skills that are not taught in regular school but they also may impact non-cognitive skills. They motivate acquisitions of work-relevant skills and provide for disadvantaged youth discipline and guidance—such as job readiness, teamwork, discipline and responsibility, leadership, flexibility, communication capacity, perseverance, and self-control. There is evidence that such programs can be effective to improve labor market prospects of vulnerable youth and women in developing countries (Card et al., 2011; Attanasio et al., 2017; Camargo et al., 2021).

However, these general policies currently have to be complemented by targeted measures to mitigate the negative effects of COVID-19 crisis on the affected youth cohort. First, they will likely not be able to complete their formal education satisfactorily because of the disruption of the regular education process. And second, the shock in the labor market made their transition more uncertain and unstable. Informality still prevails and the new technology changes the skill requirements for jobs and tasks.

4 Outlook

Every cohort of young individuals faces the recurrent challenge of the transition from school to work. This transition is influenced by the institutional arrangements at place and the prevailing economic conditions. If supportive and inclusive, labor market institutions can shorten the phase of transition and increase the chance of access to a regular job, while a heavily segmented labor market tends to keep young people in longer sequences of vulnerable, often temporary jobs.

The advent of the COVID-19 pandemic affected the current youth cohort in a particular way. On one hand, it affected the well-functioning of the education system and, on the other hand, it was a negative shock on the labor market. If it can have lasting effects on young people in the developed world, it can be more acute among youth in developing countries where the education system is in general more deficient or uneven and the labor market is more volatile and segmented.

The common policy options range from regulations that constraint the school-work choice set to a variety of policies that expand this set of choices. Among the policies that impose constraints are the minimum employment age laws and the compulsory schooling laws. Among policies that increase the choice set are measures aimed to improve the quality of formal education and policies that expand vocational and professional education. The emphasis should be placed on the second one. However, these general policies have to be complemented by targeted measures to mitigate the negative effects of COVID-19 crisis on the affected youth cohort.

References

Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *American Economic Review, 96*(3), 847–862.

Attanasio, O., Guarn, A., Medina, C., & Meghir, C. (2017). Vocational Training for Disadvantaged Youth in Colombia: A Long-Term Follow-Up. *American Economic Journal: Applied Economics, 9*(2), 131–143.

Cahuc, P., Carcillo, S., Rinne, U., & Zimmermann, K. F. (2013). Youth Unemployment in Old Europe: The Polar Cases of France and Germany. *IZA Journal of European Labor Studies, 2*, 18.

Camargo, J., Lima, L. S. E., Riva, F., & Souza, A. P. (2021). Technical Education, Non-cognitive Skills and Labor Market Outcomes: Experimental Evidence from Brazil. *IZA Journal of Labor Economics, 10*(1), 1–34.

Card, D., Ibarrarán, P., Regalia, F., Rosas-Shady, D., & Soares, Y. (2011). The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics, 29*(2), 267–300.

Cockx, B. (2016). Do Youths Graduating in a Recession Incur Permanent Losses? *IZA World of Labor, 281*. https://doi.org/10.15185/izawol.281

Edmonds, E., & Shresta, M. (2012). The Impact of Minimum Age of Employment Regulation on Child Labor and Schooling. *IZA Journal of Labor Policy, 1*, 1–14.

Eichhorst, W., & Marx, P. (2021). How Stable Is Labour Market Dualism? Reforms of Employment Protection in Nine European Countries. *European Journal of Industrial Relations, 27*(1), 93–110.

Eichhorst, W., & Rinne, U. (2017). The European Youth Guarantee: A Preliminary Assessment and Broader Conceptual Implications, IZA Policy Paper 128.

Eichhorst, W., Rodríguez-Planas, N., Schmidl, R., & Zimmermann, K. F. (2015). A Road Map to Vocational Education and Training in Industrialized Countries. *ILR Review, 68*(2), 314–337.

Emerson, P. M., & Souza, A. P. (2011). Is Child Labor Harmful? The Impact of Working Earlier in Life on Adult Earnings. *Economic Development and Cultural Change, 59*(2), 345–385.

Escudero, V., & Mourelo, E. L. (2015). The Youth Guarantee Programme in Europe: Features, Implementation and Challenges. Research Department Working Paper No. 4, International Labour Office (ILO).

ILO. (2021). An Update on the Youth Labour Market Impact of the COVID-19 Crisis.

OECD. (2020). OECD Employment Outlook.

OECD. (2021). OECD Employment Outlook.

Pastore, F. (2018). Why Is Youth Unemployment So High and Different Across Countries? *IZA World of Labor*, 420. https://doi.org/10.15185/izawol.420

A Vulnerability Approach to the Definition of the Middle Class

Luis Felipe López-Calva and Eduardo Ortiz-Juarez

1 Introduction

The concept of social class, and more specifically middle class, has been mainly discussed in sociology, attracting less attention in economics. Recently, the issue has been brought back in the discussions due to the potential growth-related and political economy implications of growing

We would like to thank Paloma Anos-Casero, Nancy Birdsall, François Bourguignon, Francisco Ferreira, Peter Lanjouw, Santiago Levy, Maria Ana Lugo, Nora Lustig, Julian Messina, Jamele Rigolini, Carlos Silva-Jauregui, and the team of the Mobility and Middle Class Regional Flagship in the office of the Chief Economist for Latin America and the Caribbean at The World Bank, as well as participants at the 15th LACEA Meeting held in Medellin, for helpful comments. The initial discussions that led to the implementation of the vulnerability approach to middle-class analysis took place while the authors were working at UNDP in New York. We are indebted to Rebeca Grynspan for her intellectual input into the shaping of this concept. Recommendations by two anonymous referees substantially improved this chapter. Megan Rounseville provided invaluable research and editing support.

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_26

middle classes in middle-income countries. Traditionally, the middle class has been defined in terms of income, consumption patterns, occupational status, or even by using self-identification. Regardless of which definition is used, the measurement of the middle class is dependent on a particular period and place and affected by factors such as asset holdings and educational levels.

Indeed, the analysis of the middle class emerges as an important issue given the strong influence of this social group on society, politics, and the economy. A vast amount of literature suggests that the middle class helps to foster economic development through its emphasis on human capital investment, consumption, and savings. This effect incentivizes a virtuous circle which contributes to its further expansion (e.g., Easterly, 2001). Other authors suggest that economic situations that are good for the middle class are also likely to be good for the poor. This fosters social cohesion and mitigates tensions between the poor and the rich; therefore, the outcome is social and political stability (e.g., Birdsall, 2010).

Over the past decade several countries in Latin America have accomplished important reductions in both poverty and inequality; as a result, policymakers are now confronting new challenges, and seeking to understand problems they face from a political economy perspective. The simultaneous emergence of improved quality and availability of data has sparked a surge in literature addressing measurement and dynamics of the middle class. However, most writing on the subject lacks clarity on the definition of what it means to be *middle class*. In the sociological literature there is a long tradition of class dynamics analysis based mainly on occupational structure following Goldthorpe's categories (Goldthorpe, 1987). In the economic literature, on the other hand, the analysis has focused mainly on *relative* definitions, addressing a stratum of the income distribution rather than an analysis of class. Most of existing relative definitions compare different middle classes from place to place because income distributions

L. F. López-Calva (✉)
Regional Bureau for Latin America and the Caribbean, United Nations Development Programme (UNDP), New York, NY, USA
e-mail: luis.lopez-calva@undp.org

E. Ortiz-Juarez
Department of International Development, King's College London, London, UK
e-mail: eduardo.ortizj@kcl.ac.uk

differ across countries.[1] An *absolute* approach thus becomes more advantageous because it identifies middle class as those households with income or consumption in a specific and comparable range. However, while most absolute measures enable comparison across countries, they have resulted in descriptive statistics of income groups rather than meaningful analysis, as they lack a conceptual framework to support a definition of thresholds.

In line with Amartya Sen's (1983) statement: "poverty is absolute in the realm of capabilities but relative in the realm of income," we propose a framework in which middle class is absolute in terms of the *functioning* that define it but relative in terms of the means through which those functioning can be achieved. Thus, just as in the case of poverty measurement, there might be arbitrariness in the selection of thresholds and the construction of income or consumption, but there is a very well-defined notion of what is being measured. We argue that vulnerability to poverty is the functioning that defines the middle class and set the lower threshold of the middle class at an absolute level. The threshold corresponds in this chapter to a 10% probability of becoming poor, based on the empirical evidence in Cruces et al. (2011a) who find that 10% of people in Latin America fell into poverty every year in a 15-year period, using synthetic panels. Though somewhat arbitrary, the relevance of this chapter lies in proposing a specific concept for middle-class measurement and making it operational, even when the thresholds to be used can vary according to the specific context.

Using a regression-based approach we exploit panel data to determine the amount of comparable income associated with that probability—using income as the relative measure of vulnerability to poverty. This creates an absolute lower threshold that is used in cross-sectional surveys to measure the middle class in Chile, Mexico, and Peru over the last two decades.

The evidence shows that both the proportion of middle-class households and the income share appropriated by this group have significantly increased in all of three countries during the period under study—in-tandem with economic growth in these countries as we should expect. It also shows that existing absolute definitions lump people who are still

[1] Recently, Foster and Wolfson (2010) proposed a threshold-free approach to measure income polarization, which can be applied to middle-class analysis. Although this approach presents a less arbitrary definition of the middle class, it is a relative measure that can tell whether the middle class is (or not) greater in two time periods, but there is no information on the magnitude of the change. A discussion and application of relative measures, related to polarization, for Latin American countries are in Cruces et al. (2011b).

poor or vulnerable to poverty into their definition of middle class, which has resulted in counter-intuitive trends—the middle class expands during economic downturns and shrinks in times of growth—a problem the vulnerability-to-poverty measure corrects. In sum, this approach makes two important contributions to the measurement of the middle class. First, conceptually, it establishes a well-defined conceptual framework for the analysis of the middle class over time. Second, methodologically, it ensures that no poor or near-poor households are being identified as middle class.

This chapter is structured as follows. Section 2 provides a brief literature review on the relationship between vulnerability and middle class. Section 3 presents an overview of existing measures of middle class and their shortcomings. Section 4 describes the data used. Finally, Sect. 5 presents the empirical strategy to estimate an absolute standard for middle-class analysis, as well as the results, and Sect. 6 concludes.

2 Linking the Notion of Vulnerability to the Definition of Middle Class

Historically, the conceptual roots of social class can be found in the work by Karl Marx and Max Weber. The Marxian approach defined class in terms of common structural positions within the organization of production, where class stratification is based on the concept of exploitation and property relations (Wright, 1979). The Weberian view, on the other hand, conceived class as individuals with common economic "life chances" which determine their opportunities for income in the market, and identified the middle class as those owning skills and education.

The latter approach has triggered an important interest in analyzing the link between class and economic vulnerability. In a recent study, Goldthorpe and McKnight (2004) show that class positions affect the risk and opportunities that individuals face. Their analysis focuses on three categories of workers and their contracts (non-skilled workers with simple contracts, professional workers and managers with comprehensive and stable contracts, and intermediate workers with "mixed" forms of contracts), as well as on three dimensions: economic security, economic stability, and economic prospects. Their main result regarding economic security is that the type of the contracts held by non-skilled workers has a direct impact on the high risks of unemployment and job loss these workers face, relative to

other classes. In this sense, the "commodifying" of labor and the associated higher probability of unemployment and less security—in terms of health and pensions—suggest that class is a central element of economic vulnerability.[2]

The relevance of analyzing vulnerability is related to the growing interest in including this concept into the design and implementation of social development strategies. Under the normative assumption that vulnerability causes economic insecurity and a reduction of welfare, it should be a central component of public actions to ensure a minimum level of economic and social protection. However, the question remains on how to measure it and how to quantify its impacts on welfare.[3] Recently, Cafiero and Vakis (2006) have suggested an approach based on an "augmented" poverty line, which in addition to including a bundle of consumption goods and services also incorporates a basic "basket of insurance" against "unacceptable risks." The main argument is that by embedding risk in the measurement of poverty, such a measure becomes a wider concept of vulnerability that captures the welfare consequences of exposure to risk, not only that of having been subject to shocks.

Given that insurance markets are actually imperfect or inexistent, households insure against risks through less efficient strategies (e.g., investing resources in activities with low returns, but which assure a minimum level of consumption), making the efficient cost of insuring against risk both variable and problematic to realistically define. We argue that it is possible to find the income level associated with a set of assets and socioeconomic characteristics that would allow the households to be less vulnerable to fall into poverty due to idiosyncratic and asymmetric shocks,[4]

[2] There is vast literature discussing and measuring the concept of vulnerability. In economics this is mainly done from the perspective of poverty traps (Ligon & Schechter, 2004; Hoddinott & Quisumbing, 2008) and poverty dynamics (Calvo & Dercon, 2007; Elbers & Gunning, 2003).

[3] The literature has proposed a number of measures designed to capture the welfare consequences of vulnerability for poor households. See for example the discussion by Dercon (2002).

[4] Longitudinal data for Chile, Mexico, and Peru—described in detail in Sect. 4—allow us to observe the incidence of events that could potentially affect welfare. In Mexico, the data show that while the incidence of loss of crops and livestock affects in higher magnitude the poorest households (1st quintile)—probably due to these activities are mostly rural, where poor population could be more exposed to shocks—an important result is that the occurrence of deaths of any household member (8.8%), economic shocks—defined as business bankruptcy or unemployment—(6.6%), and health shocks for which an individual was hospi-

and interpret it as an "augmented" poverty line in the spirit of Cafiero and Vakis. From that perspective and based on the notion discussed in Goldthorpe and McKnight, we explore the link between income and vulnerability to poverty to define the lower threshold of middle class for three Latin American countries. The conceptual basis follows the vulnerability-to-poverty approach, by which, the middle class is defined the level of income that allows individuals to protect themselves from falling into poverty over time.

3 Measurement of the Middle Class

Existing income or consumption thresholds used to measure and analyze trends for the middle class have been defined using ad-hoc criteria. Milanovic and Yitzhaki (2002) have defined the middle class as those individuals living with a per capita income on $12–50 a day, at PPP (2005 international $)[5]—roughly the mean per capita incomes of Brazil and Italy, respectively—while Kharas and Gertz (2010) have used a range of $10 and $100 daily expenditure per person, obtained after adjusting survey distributions to national accounts. These thresholds were defined by excluding those individuals who would be considered rich in Portugal and poor in Luxembourg (the poorest and richest among the industrialized countries, respectively).

Moving to a developing world perspective, Birdsall (2010) defined the middle class to include those with income between $10 a day and at or below the 95th percentile of the income distribution. She argues that $10 a day is a high income compared to the global poverty line of $1.25 a day and, although it is still low by OECD standards, it is close to setting a standard for a minimum income threshold. In an influential work, Banerjee and Duflo (2008) (BD from here onwards) have defined middle class as those living with a per capita expenditure on $2–10 a day, while Ravallion (2010) proposed thresholds ranging between $2 a day (the median

talized (12%) is somewhat stable through the income distribution suggesting thus that the entire population is prone to negative shocks. While the occurrence of health shocks is stable for all five quintiles in Mexico, in Chile and Peru it shows a growing pattern through the income distribution. Although this pattern suggests a higher use of hospitalization for upper quintiles according to the capacity to pay, these results evidence a generalized exposure to risks in all three countries.

[5] Unless otherwise indicated, all income or expenditure figures are expressed in PPP terms (2005 international $).

poverty line in 70 developing countries) and $13 a day (the poverty line in the US). Both Ravallion and BD's measures have triggered an important debate on internationally comparable measurement of the middle class.

Applying the thresholds defined by Ravallion and BD to cross-sectional data for Chile, Mexico, and Peru we observe that the middle class has decreased in the last two decades, a counterintuitive result given the rate of growth and reduction of inequality in these countries, mainly during the last decade. Conversely, and also challenging intuition, the middle class has increased in times of economic downturns moving in parallel with poverty levels.

Under these definitions, as Fig. 1 shows, the size of the middle class would have increased in Mexico during the *tequila* crisis of 1995, a period in which there was a rise in the incidence of poverty from 43.3% of households in 1994 to 60.1% in 1996. The recent global crisis of 2008–2009 caused the Mexican economy to fall by 6.3% (the GDP per capita shrank by 7.4%) and the incidence of poverty increased from 40.5% of households in 2008 to 43.7% in 2010; however, the size of the middle class would have increased again according to these definitions.

In Peru, the economy suffered a triple blow during 1998–1999 with President Fujimori's political collapse, the financial turmoil in BRIC's countries, and a serious damage caused by *El Niño* phenomenon. As a result, the incidence of poverty increased from 35.4% of households in 1998 to 39.8% in 2000; counterintuitively, the middle class would have also increased. In the following years, in spite of Peru's economic performance being among the best in Latin America since 2002, the unemployment rates were still high and moderate poverty slightly increased after a slow drop in 2001; again, the size of the middle class moves contrary to the expected. Overall, these figures show that, under those definitions, the middle class has followed a similar pattern than poverty (Fig. 2), which is indeed counter intuitive. The increase of the middle class in Mexico and Peru during economic crises relates to the low thresholds ($2–13 a day, for example) for middle class established under such definitions. Thus people who moved back into being "near poor" but were previously beyond the established thresholds would be seen as moving into the middle class. As we argue, the lack of a conceptual framework to support a definition of thresholds explains the unexpected results shown above.[6]

[6] Atkinson and Brandolini (2011) have a consistent view in the sense that they suggest that income distribution and income measures alone can no longer represent accurate definitions

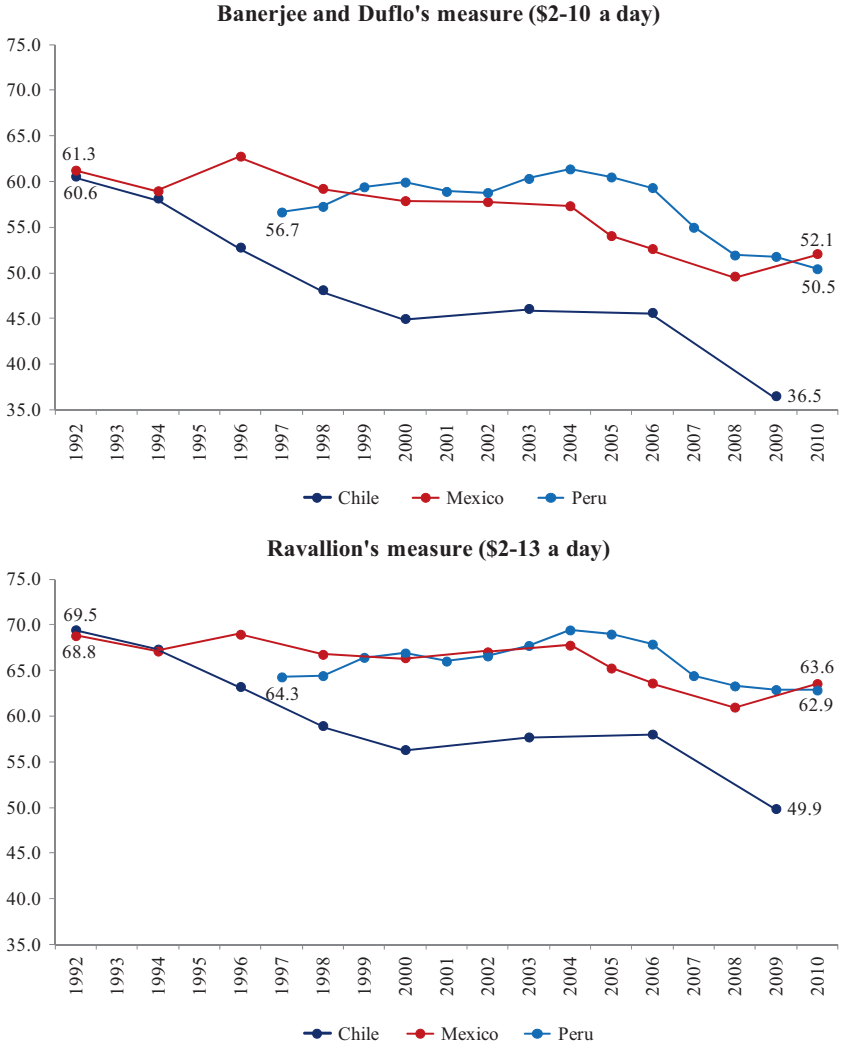

Fig. 1 People living on $2–10, and $2–13 a day in Chile, Mexico, and Peru (*Percentage of households*). (Source: Author's calculations based on cross-sectional data from CASEN, ENIGH, and ENAHO)

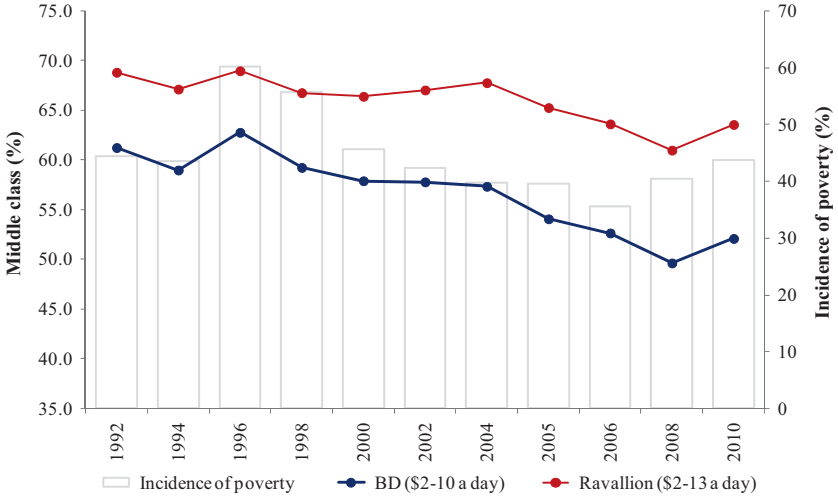

Fig. 2 People living on $2–10 and $2–13 a day; and incidence of poverty in Mexico[a] (*Percentage of households*). (Source: Author's calculations based on cross-sectional data from ENIGH. Note: [a]The left axis draws the size of the middle class according to BD and Ravallion measures, while the incidence of poverty is drawn in the right axis)

4 The Data

The vulnerability-to-poverty approach to defining the middle class is constructed in three stages. The first stage identifies actual characteristics associated with movements in or out of poverty. Then, the second and third stages construct probabilities of falling into poverty, and an income level associated to those, respectively. Based on these models, the methodology yields a monetary estimate of the lower threshold in PPP terms.

In these stages we exploit longitudinal data for Chile (2001–2006), Mexico (2002–2005), and Peru (2002–2006). In the case of Chile, we use the panel version of the Socioeconomic Characterization Survey (CASEN Panel) carried out by the Ministry of Planning (Mideplan), the Foundation for Poverty Reduction (FSP), and the Social Observatory at

of the middle class. They argue for the inclusion of wealth holdings and position in the labor market into income definitions of class. The latter has been the longest tradition in sociological work.

Universidad Alberto Hurtado (OSUAH). The first wave was undertaken in 1996 using a random sub-sample of 5210 households from the cross-sectional CASEN survey covering four regions of the country, which represent 60% of the national population. Approximately 3795 households in the baseline were re-interviewed in the wave of 2001, and 3126 of these were surveyed in 2006. These figures imply an approximately 18% attrition rate between 2001 and 2006, excluding all new households that branched out from the 2001 original ones.

In Mexico, the data is taken from the Mexican Family Life Survey (MxFLS) for the first (2002) and second (2005) rounds. The survey is representative at the national, regional, urban, and rural levels. The first wave includes approximately 8440 households distributed in 150 localities, and the second wave includes 7572 of the original households for an attrition rate of only 10%; however, only 6129 of these reported income in both waves. Finally, in Peru we use the panel of the National Household Survey (ENAHO Panel) collected by the Institute of Statistics (INEI), providing a five-wave unbalanced panel for the period 2002–2006. The size of the sample (6257 in 2002; 4217 in 2003; 6490 in 2004; 6778 in 2005; and 6593 in 2006) is representative at the national, urban, rural, and regional levels (Coast, *Sierra* and *Selva* regions). The unbalanced nature of the panel results in observations gaps since there is no information for all five waves for every household. Therefore, we use the information of the 5092 households interviewed in both 2002 and 2006 (3797 of these reported income in both waves), which imply an attrition rate of approximately 20% similar to that found in the case of Chile, and comparable to longitudinal surveys in other regions of the world. An average profile of households—from panel datasets—using most of the variables included in the analysis is presented in Table 1.

The resulting income, used as lower threshold in the definition of the middle class, is then applied to cross-sectional data in order to measure the size of this social group over the last two decades. For Chile (1992–2009), the data is taken from the cross-sectional CASEN, carried out every 2 years since 1985 by Mideplan. This data source is a nationally and regionally representative household survey used to measure the incidence of poverty and the level of inequality in the country. For Mexico (1992–2010), we use the National Household Income and Expenditure Survey (ENIGH), undertaken by the Institute of Statistics and Geography (INEGI). The ENIGH is a nationally representative survey, covering urban and rural areas, and it contains detailed information on income and

Table 1 Profile of households from longitudinal data; Chile, Mexico, and Peru (*percentage of households and averages*)

	Chile	*Mexico*	*Peru*
Education of household head in the initial year[a]			
No education	6.5%	12.7%	8.9%
Incomplete primary education	37.7%	27.5%	25.0%
Complete primary education	12.6%	21.0%	17.4%
At least 1 year of secondary education	36.0%	28.7%	34.0%
At least 1 year of tertiary education	7.2%	10.1%	14.7%
Observations	*3037*	*6108*	*3779*
Sector of activity of the household head in the initial year			
Agriculture, forestry, and fishing	22.6%	25.8%	46.1%
Mining and quarrying/electricity, gas, and water	3.0%	1.4%	1.8%
Manufacturing	15.7%	12.7%	7.2%
Construction	11.3%	11.2%	4.7%
Wholesale and retail, restaurants and hotels	16.4%	18.7%	17.7%
Transport, storage, and communications	8.1%	5.3%	6.0%
Services (private and public)	22.9%	24.9%	16.5%
Observations	*2027*	*4549*	*3258*
Marital status of the household head in the initial year			
Married	62.1%	68.8%	48.1%
Cohabiting	10.2%	12.4%	26.4%
Without partner	27.6%	18.8%	25.5%
Observations	*3126*	*6129*	*3795*
Other characteristics in the initial year			
Age of the household head (in years)	51.2	46.3	47.9
Household head male	75.9%	82.6%	81.3%
Household head with social security[b]	84.3%	57.4%	69.9%
Dirt floor in dwelling	5.8%	12.6%	45.2%
Dwelling without sanitation	4.0%	6.2%	22.0%
Rural	20.7%	51.9%	54.6%
Incidence of health shocks[c]	29.2%	12.1%	12.4%
Incidence of shocks (total)[d]		27.0%	
Observations	*3126*	*6129*	*3797*
Household per capita income, PPP (2005 international $)			
Initial year	$380.4	$353.5	$256.4
Final year (v)[e]	$384.4	$341.4	$281.9
Observations	*3126*	*6129*	*3797*

Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO

[a]The initial year is 2001 in Chile, and 2002 in Mexico and Peru

[b]In Chile, it refers to the coverage of different provisional systems (public A-D, FF.AA., ISAPRE, and private); in Mexico, it includes medical services provided by IMSS, ISSSTE, PEMEX, SEDENA, SEMAR, local institutions, universities, and private suppliers; and in Peru, it includes medical services provided by ESSALUD, *Entidad prestadora de salud*, FF.AA., *Seguro Integral de Salud*, universities, and private suppliers

[c]It refers to health shocks that demand substantial care and expenses—during 2001–2006 in Chile, 2002–2005 in Mexico, and 2002–2006 in Peru

[d]In Mexico, it includes death, illness, bankruptcy, or unemployment of any household member, and the loss of housing, business, and crop due to climate-related events

[e]The final year is 2006 in Chile and Peru, and 2005 in Mexico

different categories of expenditures. This survey is available for 1984 and 1989, and for every 2 years since 1992. Finally, in the case of Peru (1997–2010) we use data from the cross-sectional ENAHO, collected on a quarterly basis by INEI since 1997. The fourth wave of the survey is nationally representative, and it is used for poverty measurement as it collects information on income and expenditure. All these surveys collect information of dwelling characteristics, and indicators related to age, sex, employment, and education of the household members.

5 A Vulnerability Approach to Middle-Class Analysis

As we mentioned before, our empirical framework to define the middle class in terms of households' vulnerability to poverty follows a methodology comprising three stages. First, we construct poverty transition matrices from panel data using the international poverty line of $4 a day.[7] These matrices show a cross-classification of poverty status in the initial time point in the rows, and the poverty status in the final time point in the columns (Table 2), so that the proportional distributions across rows provide

Table 2 Poverty transition matrices (*row percent distribution of households*)

Chile		**2006**		
2001		Non-poor	Poor	Total
	Non-poor	93.6	6.4	100
	Poor	65.6	34.4	100
Mexico		**2005**		
2002		Non-poor	Poor	Total
	Non-poor	82.2	17.8	100
	Poor	47.4	52.6	100
Peru		**2006**		
2002		Non-poor	Poor	Total
	Non-poor	85.2	14.8	100
	Poor	37.8	62.2	100

Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO

[7] According to the World Bank, when comparing poverty rates across countries within the Latin American and the Caribbean region, the $4 a day poverty line provides a more meaningful standard compared to the $1.25 and $2 a day poverty lines which are typically used to measure poverty in less developed regions (the poverty line is a conditional mean of the national poverty lines—conditional on GDP per capita).

an assessment of the transition to poverty and the extent of poverty stability.

This analysis allows us to classify households into four categories: (1) *never poor*, if a household has never fallen under the poverty line during the 2 years analyzed in each panel; (2) *always poor*, if it has been poor in both waves of the panel; (3) *out of poverty*, if it was poor in the initial year, but exited poverty in the final year; and (4) *entered poverty*, if it was non-poor in the initial year but fell into poverty in the final one.

The idea behind calculating the actual transitions is that rather than looking at actual income we interpret it as a summary metric that gauges the level of household's asset ownership, and construct measures of permanent income by regressing income on assets and several socioeconomic characteristics. We argue that households' ability to cope with shocks is as much dependent on their actual income, as on wealth and their ability to cope with risks. Even though looking at transitions using permanent income appears to be more appropriate, because of data constraints, the middle-class threshold must ultimately use income as its metric.

In a second stage, we move to a logistic model in order to analyze the correlates of the probability of being poor over the analyzed period in each country. For a household i, the estimated probability (p_{it}) of falling into poverty between the initial (t_0) and final year (t_1) is given by:

$$p_{it} = E\left(poor_{it+1} | \mathbf{X}_{it}\right) = F\left(\mathbf{X}_{it} \bullet \boldsymbol{\beta}_t\right) \tag{1}$$

where $poor_{it+1}$ is the dependent variable taking the value of 1 if households are identified as falling into poverty between t_0 and t_1 (accounting for 11.3%, 29%, and 34% of the sample in Chile, Mexico, and Peru, respectively) and 0 otherwise; $\mathbf{X}_{it}$ is a vector of observable characteristics including demographic indicators, labor market resources, and shocks affecting the household; and $\boldsymbol{\beta}_t$ is a vector of the model parameters.[8] The vector of observable characteristics consider as demographic indicators, in t_0, rural residence, age, sex, and marital status of the household head. The last

[8] In the estimation shown below, the dependent variable includes those households in the categories *always poor* (in t_0 and t_1) and *entered poverty* (in t_1), as we are interested also in those who "remain poor." Sensitivity analysis was carried out using both definitions and the results are consistent, though the quality of the estimation is weakened due to the number of actual transitions.

variable identifies if the head is married, cohabiting, or single—including for this if she/he is widowed, separated, or never married.

For labor market resources, in t_0, the vector includes head's education level and occupational status. The head's education level is a proxy for human capital. In Chile and Peru, we use an ordinal variable with seven categories: no formal education; incomplete and complete primary; incomplete and complete secondary; and incomplete and complete tertiary education. In Mexico, we distinguish six categories: no formal education; incomplete and complete primary; complete lower-secondary; complete upper-secondary; and some year of tertiary education or more. Regarding occupational status of the head, in Chile, we use a collapsed six-class version of the EGP class classification (Erikson et al., 1979): professional and managers, clerical workers, self-employed, skilled manual workers, non-skilled manual workers, and agricultural workers. In Mexico, we use two additional categories: workers engaged in commerce and sales and army and police officers. The classification used in Peru considers those workers in agriculture, energy, manufacturing, construction, commerce, communications, and clerical activities. The model also accounts for changes between t_0 and t_1 in the number of household members engaged in paid employment, and changes in the household size.

Finally, for economic shocks affecting households a self-reported indicator is used. In Chile and Peru it includes the incidence of health shocks requiring hospitalization between t_0 and t_1, while in Mexico the variable used includes death, illness, or accident of any household member, the economic failure of a household member (unemployment and bankruptcy of business), and the loss of housing, business, crop, and livestock due to climate-related events.

In the third stage, we use the same independent variables in Eq. (1) to estimate an income equation expressed as:

$$\ln Y_{it_0} = \alpha + \mathbf{X}_{it} \bullet \boldsymbol{\beta}_t + \varepsilon_i \tag{2}$$

where $\ln Y_{it_0}$ is the household per capita income in logarithmic scale at the initial time point. Then, we calculate the average of the independent variables for an array of estimated probabilities of falling into poverty. The resulting coefficients from Eq. (2) are thus used to produce the predicted income associated with each probability. It is important to use the predicted income—a mean, conditional on characteristics—instead of the observed average income, because the predicted income has lower

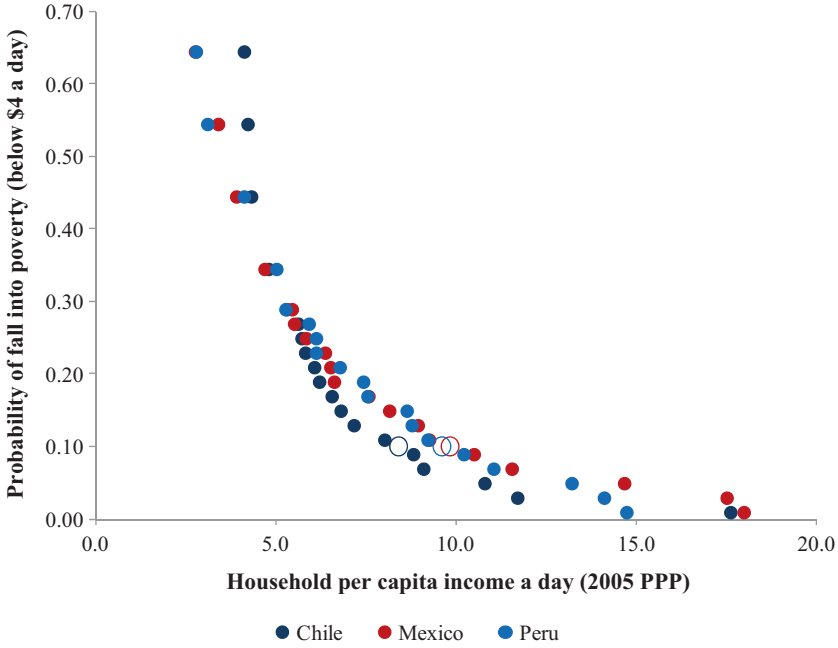

Fig. 3 Daily income by probabilities of falling into poverty; Chile, Mexico, and Peru (*household per capita income, PPP-2005 international $*). (Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO)

volatility and becomes an index related to stocks (assets), as an income generation capacity of the households.

Figure 3 shows the relationship between income and the probability that households with those levels of predicted income would fall into poverty at the end of the interval provided in each panel dataset. As the middle class, ideally, should consist of those households facing a low risk of falling into poverty over time we use a 10% probability of falling into poverty as a dividing line between economic security and vulnerability, and define the predicted income associated to that probability as the lower threshold that depicts the lower bound of the middle class.[9] Although the

[9] The lower threshold lies between the 40th (Chile) and the 60th (Mexico and Peru) percentiles of the income distribution, and it is sensitive to different probabilities of falling into poverty. In Chile, it lies in the 22th, 27th, 38th, and 52th percentiles for a probability of

selection of a probability of 0.10 is based on the empirical regularity of the share of population falling into poverty in a 15-year period using synthetic panels (Cruces et al., 2011a), it is indeed somewhat arbitrary; however, it is derived from a well-defined concept of economic security, which can be made operational for specific contexts. This concept fills a void in the literature where measures have the simple notion of "not being poor" as a middle-class concept or, even less clear, contain a tautological notion of those who have specific consumption patterns, thus resulting in either counterintuitive trends or results whose meaning is limited.

The resulting per capita incomes at a probability of 0.10 look quite similar for all three countries: non-poor individuals with a 10% probability of falling into poverty have income levels at $8.4 a day in Chile, $9.8 in Mexico, and $9.6 a day in Peru.[10] In order to not classify as middle class a number of lower class households if any of these incomes is used, we establish for strict comparability purposes a more demanding criterion of $10 a day.[11] These figures result from the complete specifications of the models shown in Appendix (Tables 6, 7, and 8), although they are robust to different model specifications (Table 3). In the case of the upper threshold we define it as $50 a day, being it an income amount that lies in the upper tail of the income distribution in all three countries.[12] Overall, we propose that the middle class is thus defined as those households living with a per capita income of $10–50 a day.

0.20, 0.15, 0.10, and 0.05, respectively. The figures in Mexico are the 39th, 49th, 58th, and 74th percentiles, respectively; while in Peru are the 40th, 52th, 58th, and 72th percentiles, respectively. The idea that looking at actual movements into poverty in the region—which are on average 10% of households every year in a 15-year period using synthetic panels (Cruces et al., 2011a)—becomes relevant. A probability to fall into poverty of, say, 30% or 50% would be inconsistent with evidence. Actually, the finding that lower levels of vulnerability are reached at relatively high percentile levels in the current distributions is a very important finding in itself for policy purposes.

[10] Estimated as the average of the independent variables for an estimated probability range between 0.09 and 0.11. Tables 6, 7, and 8 in the Appendix show the coefficients and standard errors from Eqns. (1) and (2) for Chile, Mexico, and Peru, respectively, while Figure 7 shows the correlation between the estimated probabilities from Eqn. (1) and predicted income from Eqn. (2).

[11] Ferreira et al. (2012) carry out a "validation" of the threshold by looking at income levels that are consistent with self-reported middle-class status, showing that $10 a day corresponds to the lower envelope of such income levels.

[12] In Chile, it lies in the 92th percentile in 2009, in the 97th percentile in Mexico (2010), and in the 98th percentile in Peru (2010).

Table 3 Robustness to different models specifications (*daily per capita income for a 10% probability of fall into poverty, PPP*)

Model specification	*Chile*	*Mexico*	*Peru*
Complete specification (see Tables 6, 7, and 8)	**$8.4**	**$9.8**	**$9.6**
Excluding shocks	$8.5	$9.9	$9.6
Excluding shocks and variables measuring changes	$8.8	$11.1	$11.0
Excluding variables measuring changes	$8.6	$10.9	$10.9
Excluding locational effects, shocks, and changes	$9.6	$11.3	$11.3
Excluding locational effects, marital status, shocks, and changes	$9.0	$11.4	$11.3
Only education, sex, age, and occupational status of heads	$9.8	$11.5	$11.8
Only education, sex, and age of heads	$9.6	$12.0	$11.7

Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO

Applying this absolute definition to cross-sectional surveys, available for longer time periods, the results show that the percentage of middle-class households has increased significantly from 28.3% to 53% in Chile (24.8 percentage points), from 28.8% to 40% in Mexico (15.2 points), and from 23% to 42.3% in Peru (19.4 points) over the period under analysis (Fig. 4). The growth of the middle class has been sustained and widespread: it has increased in countries recovering from crises (Mexico and Peru) and, during the global recession in 2009, it continued on its upward trend in Chile and Peru. Notably, it has increased both in fast growing countries (Chile and Peru) and slow growing ones (Mexico).

The definition of the upper threshold is not the focus of this chapter. It is worth, however, carrying out sensitive analysis for both thresholds. Moving the upper threshold up or down in the income distribution has a less relevant impact on the percentage of households in the middle class, and even less on its change. Starting from $50 a day, the decrease/increase of the size of the middle class is not larger than 5 percentage points when we apply an upper threshold ranging between $35 and $100 a day in Chile, and between $30 and $100 a day in the other two countries—although varying the upper threshold from $50 to $20 a day, it has a relatively larger impact: a reduction of 21 percentage points in Chile, and close to 14 points in Mexico and Peru (Table 4). However, if we look at the changes in the size of the middle class, variations are even smaller. Lowering the threshold from $50 to $30 per day does not affect the change in the middle class by more than 3 percentage points, on average, during the analyzed period in each country (Fig. 5). Moving the upper

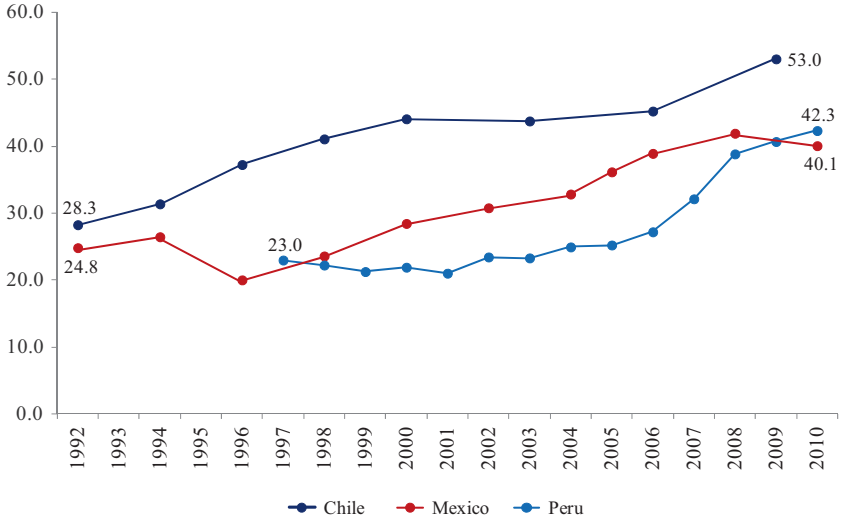

Fig. 4 Size of the middle class in Chile, Mexico, and Peru (*percentage of households*). (Source: Author's calculations based on cross-sectional data from CASEN, ENIGH, and ENAHO)

Table 4 The effect of moving the upper and lower thresholds; Chile (2009), and Mexico and Peru (2010) (*percentage points of change in the size of the middle class*)

	Moving the upper-threshold				*Moving the lower-threshold*		
	Chile	*Mexico*	*Peru*		*Chile*	*Mexico*	*Peru*
$10–20	-21.0	-13.8	-14.1	0.03–$50	-16.8	-22.9	-17.7
$10–30	-8.7	-5.0	-4.9	0.05–$50	-12.6	-16.5	-14.8
$10–35	-4.9	-3.3	-2.8	0.07–$50	-3.9	-7.3	-6.9
$10–40	-3.1	-1.7	-1.6	0.09–$50	-2.2	-3.0	-3.1
$10–50	0.0	0.0	0.0	0.10–$50	0.0	0.0	0.0
$10–60	2.1	1.1	0.9	0.13–$50	7.2	4.3	4.4
$10–70	3.4	1.8	1.4	0.15–$50	9.0	9.1	5.2
$10–80	4.4	2.1	1.7	0.17–$50	10.7	12.5	11.5
$10–90	5.0	2.4	1.9	0.19–$50	12.5	18.9	12.3
$10–100	5.2	2.6	2.0	0.21–$50	13.2	19.5	16.4

Source: Author's calculations based on cross-sectional data from CASEN, ENIGH, and ENAHO

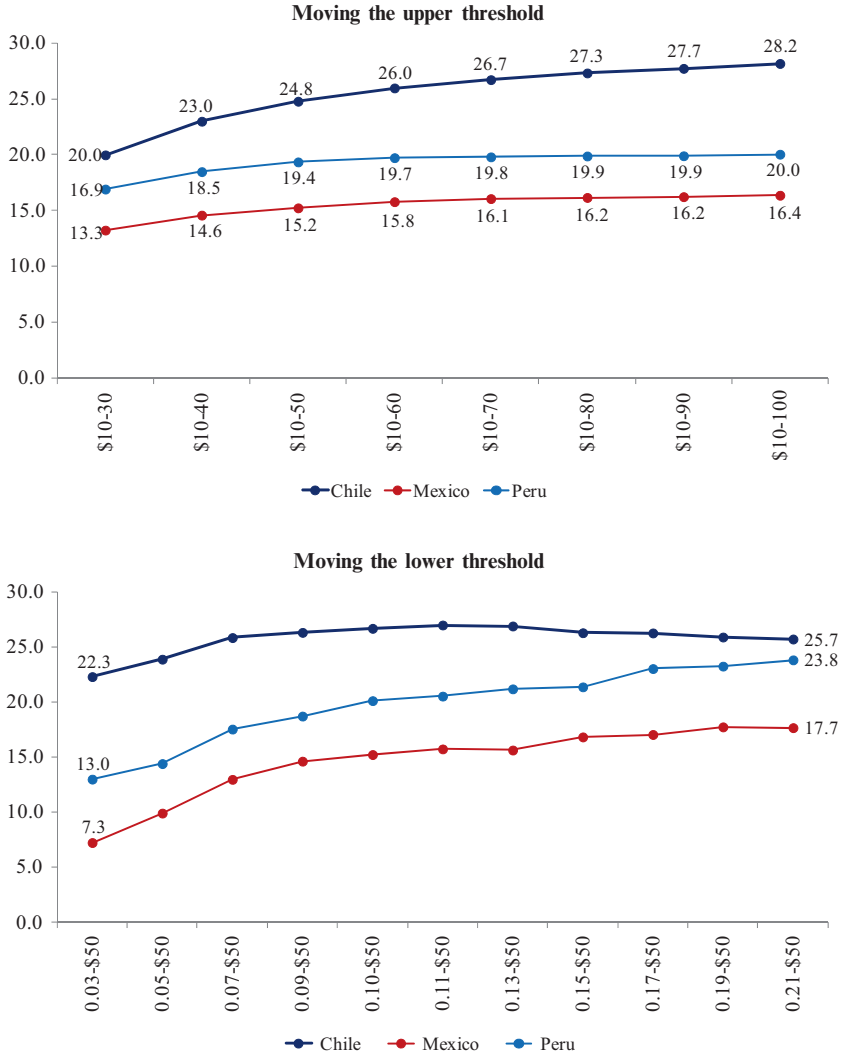

Fig. 5 The effect of moving the upper and lower thresholds; Chile (1992–2009), Mexico (1992–2010), and Peru (1997–2010) (*percentage points of change in the size of the middle class*). (Source: Author's calculations based on cross-sectional data from CASEN, ENIGH, and ENAHO)

threshold from $50 to $100 would move the percentage of Latin America's elite from the top 2% to the top 0.5%.

In contrast, varying the lower threshold has a larger impact on the size of the middle class. The sensitive analysis shows that, compared to the proposed lower threshold of 0.10, moving it from 0.03 to 0.20 would decrease/increase the size of the middle class by roughly 19 to 16 percentage points, on average, in all three countries (Table 4), and it would also affect the magnitude of changes over time, mainly in Mexico and Peru (Fig. 5). This is absolutely understandable, given the relevance of the initial conditions. Alike poverty measurement, the issue of where the poverty line falls with respect to the initial distribution matters for determining poverty elasticity to growth (Bourguignon, 2002). That is why the concept behind the notion of middle class based on vulnerability becomes even more relevant. Making such concept operational requires defining thresholds that are relevant for the specific context, while being consistent with the framework that supports the analysis. In welfare measurement, absolute concepts are always made operational with arbitrariness in one specific dimension, that is what makes the conceptual framework more relevant.

Overall, moving both the upper and lower thresholds would change the size of the middle class in a given year (Table 4); however, the main conclusion remains intact: the middle class has unambiguously increased between the initial and final points of the analyzed time span in Chile, Mexico, and Peru (Fig. 5). The middle class has not only increased in size, but it has also increased the income share appropriated by this group. In Chile, this share increased by 13.1 percentage points from 42.6% in 1992 to 55.9% in 2009; in Mexico it increased from 45.1% to 55.1% over 1992–2010, and in Peru it expanded by 16 points from 42.6% to 58.6% over 1997–2010. These general results become clear when looking at the income distribution. Figure 6 shows that the bulge in the middle class increased over the analyzed period making the bell taller and confirming the increase of the middle class, partly due to improvements in income distribution (motivated by a fall in the premium to skills and more progressive government transfers[13]), and a reasonable economic performance—at least in Chile and Peru—which is evident by the movement of the bells to the right.

[13] See, for example, Lopez-Calva and Lustig (2010) and Lustig et al. (2012).

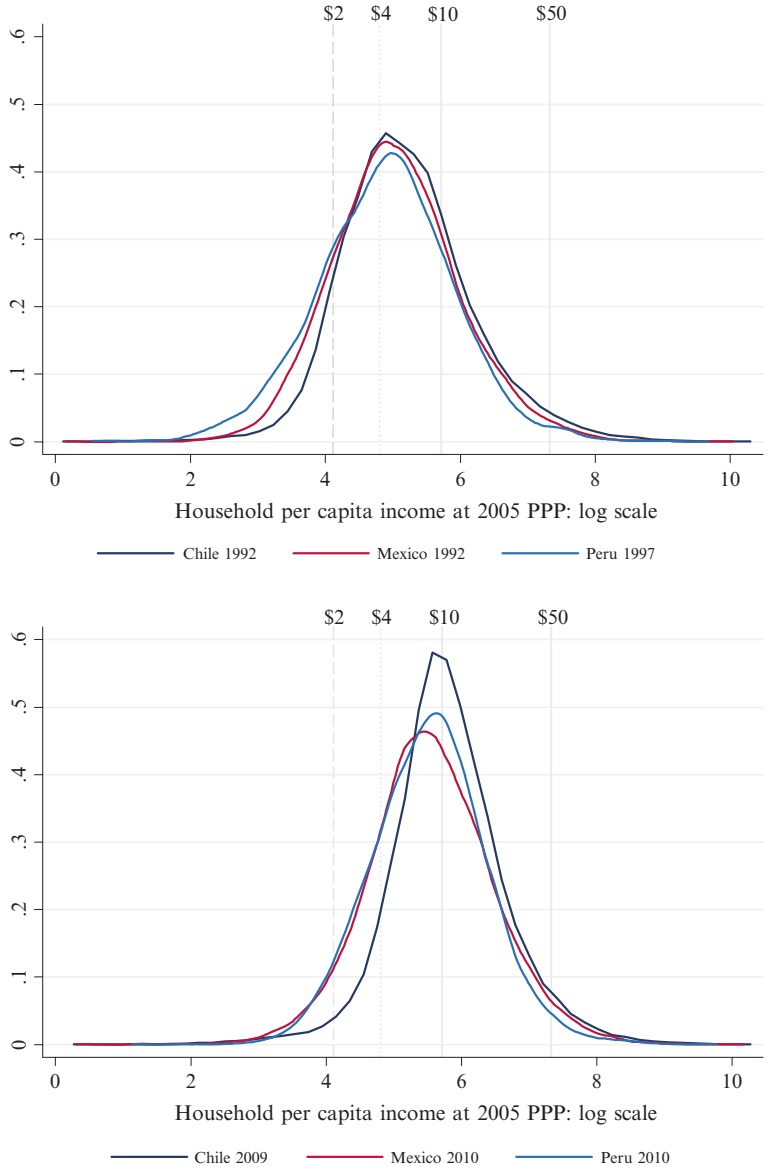

Fig. 6 Kernel distributions of income in Chile, Mexico, and Peru. (Source: Author's calculations based on cross-sectional data from CASEN, ENIGH, and ENAHO. Notes: The dashed gray line is the lower threshold of $2 a day proposed by BD and Ravallion; the dotted line is the international poverty line of $4 a day; and the solid lines are the thresholds proposed in this chapter)

In order to observe how the socioeconomic characteristics of the middle class compare relative to other groups, we estimate a two-group mean-comparison t-test to analyze if there is a significant difference between the characteristics of the middle class and, for example, those of the lower class. For that purpose, we first divide the lower class in two groups: (i) poor households, defined as those with per capita income below the international poverty line of $4 a day, and (ii) vulnerable households, defined as those with per capita income between the poverty line and $10 a day. Then, we compare both groups of households with those in middle class living on $10–50 a day. As an attempt to define a standard profile of the middle class in all three countries, Table 9 in the Appendix shows that the middle class is characterized by having higher levels of education—especially tertiary—living in urban areas, and its heads perform clerical or professional activities.

Differences between the middle class and the other groups are large and statistically significant. The four variables that show the largest differences among the group means are percentage rate of households with dirt floors, the head of the household occupation is a farmer, the head has at least 1 year of tertiary education, and rural residence (Table 5). In each case the poor

Table 5 Selected statistically significant mean comparisons[a] (*percentage of households*)

	Poor	*Vulnerable*	*Middle class*
	Chile		
Dirt floor in dwelling	10.7	6.4	3.2
Head is farmer	40.3	23.4	12.9
Rural	33.8	22.7	13.5
At least 1 year of tertiary education	1.3	3.5	11.6
	Mexico		
Dirt floor in dwelling	24.7	10.0	3.4
Head is farmer	48.4	19.5	10.3
Rural	76.8	48.5	31.5
At least 1 year of tertiary education	2.3	6.9	20.6
	Peru		
Dirt floor in dwelling	69.4	34.2	17.4
Head is farmer	71.6	31.5	16.6
Rural	77.5	44.8	26.8
At least 1 year of tertiary education	2.8	16.2	34.9

Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO

[a]A complete profile including t-tests for several socioeconomic characteristics is shown in Table 9

households have the worst prevalence rates followed by the vulnerable households and lastly by the middle class. The considerable and statistically different differences between the vulnerable and the middle class emphasize the importance of making a distinction between both groups.

In general, the evidence found suggests a higher number of middle-class households in all three countries facing a low probability of falling into poverty in the late-2000s than in the 1990s, which is undoubtedly very important in light of the reasons outlined in a large literature arguing that size and composition of the middle class are critical for strengthening and stabilizing the democratic system and its institutions, for a better economy's performance, and for social cohesion. These results differ from those obtained by applying the Ravallion and BD methods (Fig. 1). The divergences are rather evident: the absolute thresholds defined by these authors ($2–13 a day, for example, in Ravallion's case) are a very low standard for middle-class comparisons across the upper-middle-income countries analyzed in this chapter. Actually, this lower threshold is half the international poverty line of $4, suggesting thus the inclusion of near-poor and vulnerable households as part of the middle class (Fig. 6). The vulnerability-to-poverty measure of middle class proposed in this chapter corrects this shortcoming.

6 Concluding Remarks

This chapter proposes an empirical methodology to analyze the middle class based on the notion of vulnerability to poverty. Following a regression-based approach this chapter exploits panel data to estimate the amount of comparable income—$10 a day, associated with a low (0.10) probability of falling into poverty—which depicts the beginning of the middle class (lower threshold). In the case of the upper threshold it is established at $50 a day. By doing so, this chapter proposes an absolute-standard definition of the middle class as those households living with a per capita income of $10–50 a day. These thresholds are then used in cross-sectional surveys to measure the size of the middle class in Chile over 1992–2009, in Mexico over 1992–2010, and in Peru over 1997–2010.

Following this definition, this chapter overcomes the two main shortcomings that emerge when comparing middle classes across countries using existing measures. First, relative measures face the problem of different income or consumption distributions from one country to another, and thus they compare different middle classes. Second, the use of very low

absolute thresholds in upper-middle-income countries like Chile, Mexico, and Peru includes a high proportion of poor households within the middle-class definition. This sheds light on the descriptive analysis that emerges from the use of monetary thresholds without a conceptual framework.

Overall, this chapter makes two important contributions to the measurement of the middle class. First, methodologically, it ensures that no lower class or poor households are being identified as middle class. Second, conceptually, it establishes a well-defined conceptual framework for the analysis of the middle class over time. The evidence suggests a higher number of middle-class households facing a low probability of falling into poverty in the late-2000s than in the 1990s, which is undoubtedly very important in light of the reasons outlined in a large literature arguing that size and composition of the middle class are critical for strengthening and stabilizing the democratic system and its institutions, for better economic performance, and for social cohesion.

Appendix

Table 6 Determinants of falling into poverty, and income; Chile: 2001–2006[a] (*logistic and linear regression analysis*)

Model:	*Logistic*		*Linear*	
Dependent variable:	*Poverty*		*Income (log-scale)*	
	Coeff.	*S.E.*	*Coeff.*	*S.E.*
Education of the head	-0.247***	(0.068)	0.143***	(0.013)
Age of the head	0.104*	(0.054)	0.006	(0.010)
Age squared of head	-0.002**	(0.001)	0.000	(0.000)
Sex of the head (1 = male)	-0.742**	(0.316)	0.320***	(0.078)
Head without social insurance[b]	0.496**	(0.167)	-0.195***	(0.042)
Unfinished floor	0.676***	(0.243)	-0.099	(0.072)
Household without sanitation	0.271	(0.308)	-0.319***	(0.072)
Head cohabiting (omitted)	–	–	–	–
Head married	0.169	(0.236)	-0.035	(0.051)
Head without partner	-0.377	(0.329)	0.442***	(0.083)
Head in agriculture (omitted)	–	–	–	–
Head as unskilled manual worker	0.005	(0.223)	-0.025	(0.051)
Head as skilled manual worker	-0.248	(0.265)	0.088	(0.058)
Head as independent worker	-0.120	(0.252)	0.158***	(0.059)

(*continued*)

Table 6 (continued)

Model:	*Logistic*		*Linear*	
Dependent variable:	*Poverty*		*Income (log-scale)*	
	Coeff.	*S.E.*	*Coeff.*	*S.E.*
Head in clerical activities	-0.463	(0.369)	0.146**	(0.066)
Head as professional manager	-2.260**	(1.069)	0.427***	(0.088)
Region VII (omitted)	–	–	–	–
Region III	-0.587*	(0.350)	0.293***	(0.074)
Region VIII	0.122	(0.182)	0.018	(0.042)
Metropolitan region	-0.704***	(0.216)	0.247***	(0.041)
Rurality	0.277	(0.189)	-0.061	(0.046)
Occurrence of health shocks 2001–2006	0.128	(0.176)	0.037	(0.038)
Change in number of members working 2001–2006	-0.490***	(0.094)	-0.158***	(0.017)
Change in household size 2001–2006	0.190*	(0.099)	0.164***	(0.022)
Constant	-1.537	1.277	9.603***	(0.248)
Observations	1950		1950	
Pseudo R^2/R^2	0.148		0.385	

Source: Author's calculations based on data from CASEN Panel 2001–2006

Robust standard errors in parentheses, *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

[a]Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model

[b]It refers to the coverage of different provisional systems (public A-D, FF.AA., ISAPRE, and private).

Table 7 Determinants of falling into poverty, and income; Mexico: 2002–2005[a] (*logistic and linear regression analysis*)

Model:	*Logistic*		*Linear*	
Dependent variable:	*Poverty*		*Income (log-scale)*	
	Coeff.	*S.E.*	*Coeff.*	*S.E.*
Education of the head	-0.328***	(0.052)	0.228***	(0.017)
Age of the head	-0.019	(0.017)	-0.005	(0.006)
Age squared of head	0.000	(0.000)	0.000*	(0.000)
Sex of the head (1 = male)	-0.364**	(0.166)	0.356***	(0.068)
Head without social insurance[b]	0.571***	(0.083)	-0.310***	(0.032)

(*continued*)

Table 7 (continued)

Model:	*Logistic*		*Linear*	
Dependent variable:	*Poverty*		*Income (log-scale)*	
	Coeff.	*S.E.*	*Coeff.*	*S.E.*
Unfinished floor	0.615***	(0.107)	-0.348***	(0.046)
Household without sanitation	0.429***	(0.137)	-0.143**	(0.067)
Head cohabiting (omitted)	–	–	–	–
Head married	-0.086	(0.112)	-0.017	(0.038)
Head without partner	-0.691***	(0.174)	0.529***	(0.067)
Head in agriculture (omitted)	–	–	–	–
Head as unskilled manual worker	-0.404***	(0.133)	0.367***	(0.050)
Head as skilled manual worker	-0.734***	(0.108)	0.383***	(0.045)
Head as independent worker	-0.522**	(0.205)	0.279***	(0.083)
Head in clerical activities	-0.793***	(0.231)	0.589***	(0.064)
Head as professional manager	-0.926***	(0.211)	0.706***	(0.065)
Head in commerce and services	-0.493***	(0.140)	0.277***	(0.058)
Head in army, police, and other	-0.796***	(0.224)	0.328***	(0.071)
South region (omitted)	–	–	–	–
Central region	-0.321***	(0.111)	0.174***	(0.043)
Western region	-0.523***	(0.118)	0.259***	(0.046)
Northwest region	-0.800***	(0.122)	0.338***	(0.042)
Northeast region	-0.523***	(0.119)	0.354***	(0.043)
Rurality	0.684***	(0.091)	-0.351***	(0.031)
Occurrence of shocks 2002–2005[c]	0.156*	(0.087)	-0.034	(0.032)
Change in number of members working 2002–2005	-0.420***	(0.038)	-0.118***	(0.013)
Change in household size 2002–2005	0.199***	(0.040)	0.034**	(0.014)
Constant	0.646	(0.497)	5.956***	(0.181)
Observations	4514		4514	
Pseudo R^2/R^2	0.196		0.355	

Source: Author's calculations based on data from MxFLS 2002–2005

Robust standard errors in parentheses, *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

[a]Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model

[b]It includes medical services provided by IMSS, ISSSTE, PEMEX, SEDENA, MARINA, local institutions, universities, and private suppliers

[c]It includes death, illness, bankruptcy, or unemployment of any household member, and the loss of housing, business, and crop due to climate-related events

Table 8 Determinants of falling into poverty, and income; Peru: 2002–2006[a] (*logistic and linear regression analysis*)

Model:	*Logistic*		*Linear*	
Dependent variable:	*Poverty*		*Income (log-scale)*	
	Coeff.	*S.E.*	*Coeff.*	*S.E.*
Education of the head	-0.485***	(0.053)	0.244***	(0.014)
Age of the head	-0.063***	(0.019)	0.012**	(0.005)
Age squared of head	0.000**	(0.000)	0.000	(0.000)
Sex of the head (1 = male)	-0.189	(0.185)	0.172***	(0.053)
Head without social insurance[b]	-0.181*	(0.100)	0.129***	(0.026)
Unfinished floor	0.786***	(0.100)	-0.262***	(0.026)
Household without sanitation	0.498***	(0.105)	-0.143***	(0.031)
Head cohabiting (omitted)	–	–	–	–
Head married	-0.216**	(0.109)	0.093***	(0.027)
Head without partner	-0.877***	(0.176)	0.404***	(0.050)
Head in agriculture (omitted)	–	–	–	–
Head in mining, electricity, gas, and water	-1.650***	(0.401)	0.971***	(0.108)
Head in manufacturing	-0.523**	(0.211)	0.358***	(0.050)
Head in construction	-0.530**	(0.231)	0.288***	(0.053)
Head in commerce	-0.650***	(0.152)	0.422***	(0.038)
Head in transport and communications	-0.583***	(0.221)	0.358***	(0.047)
Head in government and clerical activities	-1.412***	(0.248)	0.543***	(0.048)
Head in other services	-0.823***	(0.221)	0.332***	(0.048)
Selva region (omitted)	–	–	–	–
North Coast	-0.753***	(0.136)	0.196***	(0.035)
Central Coast	-1.458***	(0.232)	0.228***	(0.046)
Southern Coast	-0.988***	(0.282)	0.308***	(0.057)
Northern Sierra	-0.079	(0.195)	-0.065	(0.061)
Central Sierra	0.016	(0.138)	-0.111***	(0.039)
Southern Sierra	0.106	(0.152)	-0.058	(0.041)
Metropolitan area of Lima	-0.192***	(0.226)	0.299***	(0.043)
Rurality	0.432***	(0.129)	-0.072**	(0.030)
Occurrence of health shocks 2002–2006	-0.570***	(0.147)	0.166***	(0.037)
Change in number of members working 2002–2006	-0.305***	(0.045)	-0.104***	(0.010)
Change in household size 2002–2006	0.236***	(0.030)	0.076***	(0.007)
Constant	3.056***	(0.546)	3.726***	(0.145)
Observations	3258		3258	
Pseudo R^2/R^2	0.272		0.489	

Source: Author's calculations based on data from ENAHO Panel 2002–2006

Robust standard errors in parentheses, *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

[a]Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model

[b]It includes medical services provided by ESSALUD, *Entidad prestadora de salud*, FF.AA., *Seguro Integral de Salud*, universities, and private suppliers

Table 9 Two-group mean-comparison *t*-test; Chile, Mexico, and Peru (*percentage of households and levels of significance*)

	Incidence by group			*Level of significance of differences between groups (ref. 95%)*	
	Poor	*Vulnerable*	*Middle class*		
Groups →	*<$4*	*$4–10*	*$10–50*	*<$4/$10–50*	*$4–10/$10–50*
Chile					
No education	4.9%	7.3%	6.6%	0.170	0.492
Incomplete primary	47.5%	38.0%	33.7%	0.000	0.025
Complete primary	18.9%	12.9%	10.3%	0.000	0.050
Complete secondary	8.8%	16.8%	16.8%	0.000	0.996
At least 1 year of tertiary	1.3%	3.5%	11.6%	0.000	0.000
With social security[a]	78.8%	84.8%	86.4%	0.000	0.271
Dirt floor in dwelling	10.7%	6.4%	3.2%	0.000	0.000
Dwelling without sanitation	8.5%	4.7%	1.2%	0.000	0.000
Head is farmer	40.3%	23.4%	12.9%	0.000	0.000
Head is an unskilled manual worker	26.8%	29.7%	21.6%	0.059	0.000
Head is a skilled manual worker	11.8%	15.7%	12.2%	0.872	0.044
Head in clerical activities	5.4%	9.6%	13.8%	0.000	0.008
Head is professional/ manager	0.8%	3.9%	13.1%	0.000	0.000
Rural	33.8%	22.7%	13.5%	0.000	0.000
Incidence of health shocks 2001–2006	24.1%	26.5%	33.7%	0.000	0.000
Mexico					
No education	20.5%	11.1%	6.9%	0.000	0.000
Incomplete primary	36.8%	27.5%	18.9%	0.000	0.000
Complete primary	22.5%	21.8%	19.0%	0.007	0.029
At least 1 year of secondary[b]	14.4%	24.6%	22.8%	0.000	0.194
At least 1 year of tertiary	2.3%	6.9%	20.6%	0.000	0.000
With social security[a]	35.2%	60.7%	76.2%	0.000	0.000
Dirt floor in dwelling	24.7%	10.0%	3.4%	0.000	0.000
Dwelling without sanitation	11.6%	5.0%	2.0%	0.000	0.000
Head is farmer	48.4%	19.5%	10.3%	0.000	0.000
Head is an unskilled manual worker	10.6%	13.1%	9.9%	0.534	0.006
Head is a skilled manual worker	20.5%	33.6%	30.8%	0.000	0.098

(*continued*)

Table 9 (continued)

	Incidence by group			*Level of significance of differences between groups (ref. 95%)*	
	Poor	*Vulnerable*	*Middle class*		
Groups →	*<$4*	*$4–10*	*$10–50*	*<$4/$10–50*	*$4–10/$10–50*
Head in clerical activities	1.2%	5.9%	9.8%	0.000	0.000
Head is professional/ manager	2.3%	5.8%	19.0%	0.000	0.000
Rural	76.8%	48.5%	31.5%	0.000	0.000
Incidence of health shocks 2002–2005	12.1%	13.2%	11.0%	0.282	0.036
Incidence of shocks (total) 2002–2005[c]	28.4%	28.3%	24.2%	0.003	0.003
Peru					
No education	13.0%	7.6%	2.7%	0.000	0.000
Incomplete primary	34.9%	21.8%	11.3%	0.000	0.000
Complete primary	21.3%	15.7%	13.3%	0.000	0.152
Complete secondary	12.1%	23.1%	25.7%	0.000	0.190
At least 1 year of tertiary	2.8%	16.2%	34.9%	0.000	0.000
With social security[a]	71.2%	67.8%	71.3%	0.948	0.100
Dirt floor in dwelling	69.4%	34.2%	17.4%	0.000	0.000
Dwelling without sanitation	35.9%	15.0%	6.9%	0.000	0.000
Head in agriculture	71.6%	31.5%	16.6%	0.000	0.000
Head in mining and energy	0.6%	2.1%	4.1%	0.000	0.012
Head in manufacturing	3.8%	9.0%	12.0%	0.000	0.052
Head in sales, restaurants, and hotels	9.5%	22.8%	26.1%	0.000	0.136
Head in public administration	2.4%	9.6%	15.0%	0.000	0.001
Rural	77.5%	44.8%	26.8%	0.000	0.000
Incidence of health shocks 2002–2006	8.7%	14.0%	16.4%	0.000	0.144

Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO

[a]In Chile, it refers to the coverage of different provisional systems (public A-D, FF.AA., ISAPRE, and private); in Mexico, it includes medical services provided by IMSS, ISSSTE, PEMEX, SEDENA, MARINA, local institutions, universities, and private suppliers; and in Peru, it includes medical services provided by ESSALUD, *Entidad prestadora de salud*, FF.AA., *Seguro Integral de Salud*, universities, and private suppliers

[b]It includes lower-secondary and upper-secondary

[c]It includes death, illness, bankruptcy, or unemployment of any household member, and the loss of housing, business, and crop due to climate-related events

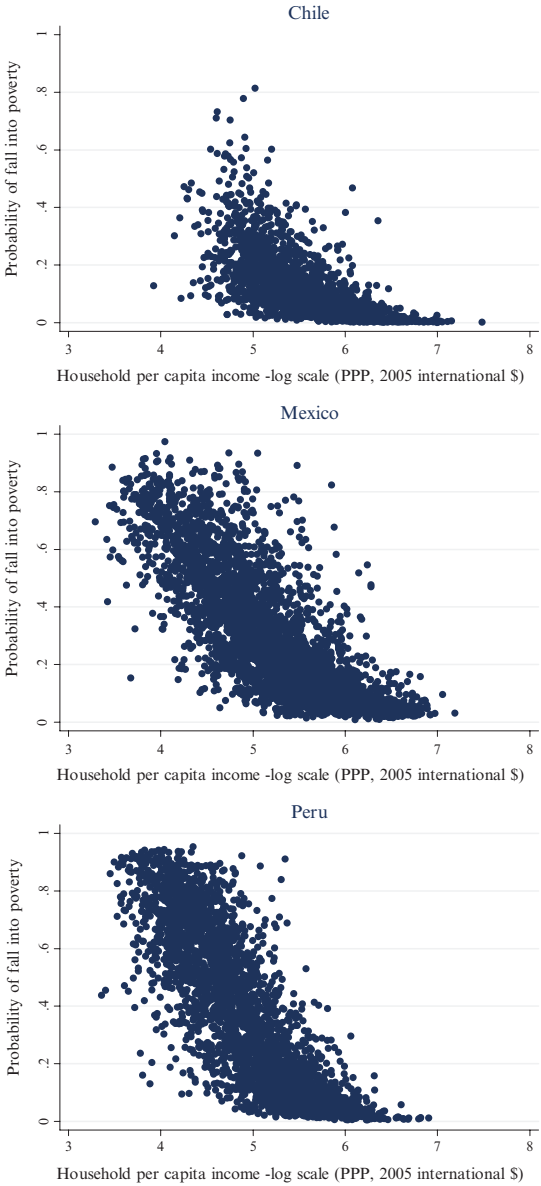

Fig. 7 Correlation between the estimated probabilities of fall into poverty and—predicted—income; Chile, Mexico, and Peru. (Source: Author's calculations based on longitudinal data from CASEN, MxFLS, and ENAHO)

References

Atkinson, A. B., & Brandolini, A. (2011). On the Identification of the 'Middle Class'. In J. C. Gornick & M. Jaentti (Eds.), *Inequality and the Status of the Middle Class*. Stanford University Press.

Banerjee, A., & Duflo, E. (2008). What is Middle Class About the Middle Classes Around the World? *Journal of Economic Perspectives, 22*(2), 3–28.

Birdsall, N. (2010). The (indispensable) Middle Class in Developing Countries. In R. Kanbur & M. Spence (Eds.), *Equity and Growth in a Globalizing World*. The World Bank.

Bourguignon, F. (2002). *The Growth Elasticity of Poverty Reduction: Explaining Heterogeneity across Countries and Time Periods*. DELTA Working Paper 2002–03. DELTA, Ecole normale supérieure.

Cafiero, C., & Vakis, R. (2006). *Risk and Vulnerability Considerations in Poverty Analysis: Recent Advances and Future Directions*. SP Discussion Paper 0610. The World Bank.

Calvo, C., & Dercon, S. (2007). *Vulnerability to poverty*. CSAE Working Paper 2007–03. University of Oxford.

Cruces, G., Lanjouw, P., Lucchetti, L., Perova, E., Vakis, R., & Viollaz, M. (2011a). *Intra-generational Mobility and Repeated Cross-sections. A Three-country Validation Exercise*. Policy Research Working Paper 5916. The World Bank.

Cruces, G., Lopez-Calva, L. F., & Battiston, D. (2011b). *Down and Out or up and in? In Search of Latin America's Elusive Middle Class*. CEDLAS Working Paper 113. Universidad Nacional de La Plata.

Dercon, S. (2002). Income Risk, Coping Strategies and Safety Nets. *World Bank Research Observer, 17*(2), 141–166.

Easterly, W. (2001). Middle Class Consensus and Economic Development. *Journal of Economic Growth, 6*(4), 317–336.

Elbers, C., & Gunning, J. (2003). *Vulnerability in a Stochastic Dynamic Model*. Discussion Paper 2003–070/2. Tinbergen Institute.

Erikson, R., Goldthorpe, J., & Portocarero, L. (1979). Intergenerational Class Mobility in Three Western European Societies. *British Journal of Sociology, 30*(4), 415–441.

Ferreira, F., Messina, J., Rigolini, J., Lopez-Calva, L. F., Lugo, M. A., & Vakis, R. (2012). *Economic Mobility and the Latin American Middle Class*. The World Bank.

Foster, J., & Wolfson, M. (2010). Polarization and the Decline of the Middle Class: Canada and the U.S. *Journal of Economic Inequality, 8*(2), 247–273.

Goldthorpe, J. H. (1987). *Social Mobility and Class Structure in Modern Britain*. Clarendon Press.

Goldthorpe, J. H., & McKnight, A. (2004). *The Economic Basis of Social Class*. CASE Paper 80. London School of Economics.

Hoddinott, J., & Quisumbing, A. (2008). *Methods for Microeconometric Risk and Vulnerability Assessments*. Available at SSRN: https://doi.org/10.2139/ssrn.1281055

Kharas, H., & Gertz, G. (2010). The New Global Middle Class: A Cross-over from West to East. In C. Li (Ed.), *China's Emerging Middle Class: beyond Economic Transformation*. Brookings Institution Press.

Ligon, E., & L. Schechter (2004). *Evaluating Different Approaches to Estimating Vulnerability*. SP Discussion Paper 0210. The World Bank.

Lopez-Calva, L. F., & Lustig, N. (Eds.). (2010). *Declining Inequality in Latin America: A Decade of Progress?* Brookings Institution Press.

Lustig, N., Lopez-Calva, L. F., & Ortiz-Juarez, E. (2012). Declining Inequality in Latin America in the 2000s: The Cases of Argentina, Brazil, and Mexico. *World Development*. https://doi.org/10.1016/j.worlddev.2012.09.013

Milanovic, B., & Yitzhaki, S. (2002). Does Decomposing World Income Distribution: Does the World Have a Middle Class? *Review of Income and Wealth, 48*(2), 155–178.

Ravallion, M. (2010). The Developing World's Bulging (but Vulnerable) Middle Class. *World Development, 38*(4), 445–454.

Sen, A. (1983). Poor, Relatively Speaking. *Oxford Economic Papers, 35*(2), 153–169.

Wright, E. O. (1979). *Class Structure and Income Determination*. Academic Press.

Trade Unions, Work and Resilience

Elizabeth Cotton and Miguel Martinez Lucio

1 The Resilience Debates

The debates within trade unions and labour networks about resilience and the use of workplace wellbeing programmes are highly politicised and critical. There is a tendency for resilience to be resisted within trade unions partly due to its strategic use by governments to link health and work outcomes using wellbeing (Cederstrom & Spicer, 2015) and resilience programmes to increase economic activity and reduce related healthcare and welfare benefit costs. This is exemplified graphically by the UK's introduction of 'talking therapies' into unemployment and welfare benefit structures, described by radical mental health groups as the introduction of a system of 'psycho-compulsion' (Freidli & Stern, 2019). Current critiques of the growing policy links between health and work challenge the assumption that work is good for our mental health (Frayne, 2019) as well

E. Cotton (✉)
Cardiff School of Management, Cardiff Metropolitan University, Cardiff, UK
e-mail: ECotton@cardiffmet.ac.uk

M. Martinez Lucio
University of Manchester, Manchester, UK
e-mail: Miguel.MartinezLucio@manchester.ac.uk

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_27

as the growing body of research that maps the impact of insecure and precarious work, including digitalization, on our states of mind.

Additionally, the use of resilience programmes to support workers undergoing restructuring or redundancies, often in relation to large-scale cuts in the public sector since 2008, has increased steadily since 1991. This use of resilience creates a tension for trade unions between engaging with genuine attempts to help workers adapt to stressors at work and challenging strategic attempts to obscure poor working conditions such as work intensification (Boxall & Macky, 2014; Carter et al., 2013) or the impact of insecure work in the gig economy (Woods et al., 2019).

Given the model of occupational stress and risk assessment that predominantly shapes trade union mental health responses (Martinez Lucio, 2020) it is no surprise then that workplace resilience initiatives have been pushed back by trade unions as an attempt to pathologise the natural responses of workers to workplace stress. This includes the rejection of a reactive and individualised set of resilience techniques rather than addressing the workplace factors that cause and compound mental health problems (HSE WHEC, 2021). Further, that a more radical model of resilience is required to address the dynamic, relational and political factors at play in determining worker resilience.

2 The Resilience Landscape 1991–Now

Two key societal developments have heavily influenced the politicisation of resilience at work. Firstly the growing public engagement with the link between social and health inequalities (Marmot, 2020) which has opened up debates within mental health about the evidenced link between discrimination, poverty and social exclusion and mental health problems. This has helped to develop an understanding within workplaces of the direct mental health consequences of workplace racism and bullying, for example, which can feed into the development of more radical models of resilience.

The second and connected development is the degrading of mental health services, many of which were already severely underfunded and inadequate in terms of their quality and their accessibility (Cotton, 2019). What has emerged in public health systems is a standardised and medical model of psychiatric disorder diagnoses and use of prescription medication, with increasingly short-term and digitalized cognitive and behavioural interventions rather than investment into psychodynamic talking

therapies and patient-centred forms of treatment (Davies, 2021). This has resulted in a split between public and private models of treatment, which further compounds the inequalities of access to quality care.

In response to this, sociological perspectives on resilience at work question the dominant model of resilience and wellbeing used in the workplace which is based on a positive psychology model, promoted through the work of Seligman (2011), using cognitive and behavioural techniques including mindfulness practices (Purser, 2019) to promote positive attitudes and behaviours. The concern with this model is that it rests on a unitarist approach and unilateral workplace programmes which individualise collective and psychosocial factors that ultimately underpin mental health problems. In so doing it represents a strategic attempt to promote individual resilience as an inadequate response to systemic organisational problems and poor working conditions.

3 Resilience as a Contested Terrain and the Abuse of the Concepts of Wellbeing and Work-Life Balance

Within this context of change and deteriorating support for mental health issues we find the concept of resilience has become a contested terrain. Trade unions and related organisations have become weary of the way employers utilise this concept in the way they respond to the ongoing and intensifying mental health challenges workers face across a range of fronts. The tendency to mobilise the language of 'wellbeing' as a set of practices that can be transferred to the workforce for them to manage their way through day-to-day difficulties is increasingly common. Part of the problem is that in our attempt to widen the debate on wellbeing and capture its essence or diversity (see Dodge et al., 2012) there is a tendency to avoid a deeper explanation of the causes of stress and mental health problems, for example, and the way contemporary workplace regimes are fundamentally contributing to it.

However, within the study of the sociology of work there has, nevertheless, been a concerted attempt to link the nature of managerial change and organisational control strategies to the declining nature of wellbeing (Stewart et al., 2009; Carter et al., 2011, 2013). To that extent, the question of resilience leads to a range of uncertainties and disquiet given that it potentially closes the discussion on questions of causation and

responsibility. The focus instead moves to the question of resolving the problems of mental health at work through advisory and minimalist, individualised corporate welfare strategies (if at all).

Nevertheless, Stewart (2006) highlighted the way such 'individual' issues were becoming part of a broader rubric of mobilisation and collective concerns within the workplace and within worker organisations: these in turn, as stated earlier, have become more politicised as the overall frameworks of support for mental health have been steadily collapsing under the pressure of funding cuts and financial austerity. What is more, trade unions have increasingly seen the question of health and safety in relation to specific issues of mental health and stress as requiring a broadening of the concept of resilience per se. Be it through individual case representation, collective representation within the structures of voice within an organisation and the mobilising and lobbying that they lead on in relation to the state and social movements on questions of mental health and stress, for example, trade unions have attempted to generate a series of counterpoints with regards to the way management attempt to individualise and fragment the question of resilience (Martinez Lucio, 2020). In fact, mental health at work has become the basis for a much wider dialogue on the purpose of work and the way it should develop in relation to work design, working time and the resources that are required for a more participative dialogue.

To this extent, the nature of resilience is increasingly being expanded and reimagined in terms of the nature of the body and the way we understand a 'balanced life': the latter context assuming that work-life balance issues should not be subsumed within the individualised understandings that management seem to be propagating. What we are seeing—perhaps by default more than by design—is a trade union response to the need to widen this concept of resilience and link it to an ergonomic and more plural understanding of the way workers and the body exist. For example, work-life balance is a concept that is highly problematic when the overall social and personal circumstances of an individual—and the lack of support they have in the non-work sphere—are not taken into account: this often leads to limits on work not being paralleled with interventions within the non-work sphere which for women, for example, usually consist of further work and toil (Hilbrecht et al., 2008).

Finally, an often-overlooked structure within trade unions that offers an alternative strategy in building worker resilience relates to their workplace education provision (Cotton, 2020). Trade union education, based on

worker education methods and heavily influenced by the emancipatory education model developed by Paulo Freire (1970), adopts a critical perspective and examines power relations, explores emotional experience and is action oriented. For this reason, issues of mental health are often raised and processed within these educational spaces, and trade union educators routinely will work with mental health, wellbeing and resilience curricula. Because this model of learning is based on collective reflection it is a methodology which prioritises building relationality and as such offers an important support to members not least in its emphasis on collectively addressing problems at work. Although this function is under-articulated within many although not all trade unions, the pedagogic model potentially provides an alternative approach to building workplace resilience from the dominant and individualistic positive psychology model and therefore represents a broader strategic trade union response to the current mental health crisis. Central to these developmental spaces is an alternative view of what it is to understand the circumstances within which we find ourselves. These spaces counter the hierarchical or pseudo-participative narratives of more corporate approaches which are ambivalent—at best—about the way experiences of work should be reflected upon.

What we are therefore noting is a seismic shift in the way resilience is being understood, with a lack of consensus as to what it means and what role it plays. There seems to be a competing set of approaches that hinge on whether you connect, or not, the need for greater resilience with the deeper structures and drivers that are driving the deterioration of working conditions. Furthermore, locating resilience across a wider set of activities—the social, the creative and the developmental—has been pivotal to trade union responses. However, the temptation is to simply accord health and safety roles a more technocratic meaning and to simply develop short-term support for workers who are at the sharper end of work intensification. For this reason, future research and discussion may have to place a greater emphasis on the purpose of work and the place it has within our overall development: questions of resilience need to be continuously widened. Hence, trade unions are being pulled by the need to service and resolve specific issues of mental health, on the one hand, and developing a broader political engagement with the way work has become dysfunctional and unbalanced.

References

Boxall, P., & Macky, K. (2014). High-involvement Work Processes, Work Intensification and Employee Well-Being. *Work, Employment and Society, 28*(6), 963–984.

Carter, B., Danford, A., Howcroft, D., Richardson, H., Smith, A., & Taylor, P. (2011). All They Lack Is a Chain: Lean and the New Performance Management in the British Civil Service. *New Technology, Work and Employment, 26*(2), 83–97.

Carter, B., Danford, A., Howcroft, D., Richardson, H., Smith, A., & Taylor, P. (2013). 'Stressed Out of My Box': Employee Experience of Lean Working and Occupational Ill-Health in Clerical Work in the UK Public Sector. *Work, Employment and Society, 27*(5), 747–767.

Cederstrom, C., & Spicer, A. (2015). *The Wellness Syndrome*. Polity Press.

Cotton, E. (2019). The Industrial Relations of Mental Health. In C. Jackson & R. Rizq (Eds.), *The Industrialisation of Care: Counselling & Psychotherapy in a Neoliberal Age*. PCCS Books.

Cotton, E. (2020). Wellbeing on the Healthcare Frontline: A Safe Laboratory for Critical Action Learning. *Academy of Management Learning & Education, 20*(4). https://doi.org/10.5465/amle.2018.0319

Davies, J. (2021). *Sedated: How Modern Capitalism Created Our Mental Health Crisis*. Atlantic Books.

Dodge, R., Daly, A. P., Huyton, J., & Sanders, L. D. (2012). The Challenge of Defining Wellbeing. *International Journal of Wellbeing, 2*(3), 222–235.

Frayne, D. (Ed.). (2019). *The Work Cure: Critical Essays on Work and Wellness*. PCCS Books.

Freidli, L., & Stern, R. (2019). Positive Affect as Coercive Strategy: Conditionality, Activation and the Role of Psychology in UK Government Workfare Programmes. In R. Rizq & C. Jackson (Eds.), *The Industrialisation of Care*. PCCS Books.

Freire, P. (1970). *Pedagogy of the Oppressed*. Continuum Books.

Hilbrecht, M., Shaw, S. M., Johnson, L. C., & Andrey, J. (2008). 'I'm Home for the Kids': Contradictory Implications for Work–Life Balance of Teleworking Mothers. *Gender, Work & Organization, 15*(5), 454–476.

HSE WHEC. (2021). Evaluating Interventions in Work-Related Ill Health and Disease. WHEC Report. Health & Safety Executive.

Layard, R. (2011). *Happiness: Lessons from a New Science*. Penguin.

Malenfant, R., Larue, A., & Vezina, M. (2007). Intermittent Work and Well-Being—One Foot in the Door. *One Foot Out. Current Sociology, 55*(6), 814–835.

Marmot, M. (2020). *Health Equality in England: The Marmot Review 10 years on*. The Institute of Health Inequality. Retrieved April 6, 2020, from https://www.health.org.uk/sites/default/files/upload/publications/2020/

Health%20Equity%20in%20England_The%20Marmot%20Review%2010%20Years%20On_full%20report.pdf

Martinez Lucio, M. (2020). Trade Unions and Stress at Work: The Evolving Responses and Politics of Health and Safety Strategies in the Case of the UK. In R. J. Burke & S. Pignata (Eds.), *Handbook of Research on Stress and Well-Being in the Public Sector* (pp. 15–32). Edward Elgar Publishing Ltd.

Purser, R. (2019). *McMindfulness: How Mindfulness Became the New Capitalist Spirituality*. Repeater Books.

Seligman, M. E. P. (2011). *Flourish: A New Understanding of Happiness and Well-Being and How To Achieve Them*. Nicholas Brearley Publishing.

Stewart, P. (2006). Individualism and Collectivism in the Sociology of the Collective Worker. In L. E. Alonso & M. Martínez Lucio (Eds.), *Employment Relations in a Changing Society: Assessing the Post-Fordist Paradigm* (pp. 182–199). Routledge.

Stewart, P., Murphy, K., Danford, A., Richardson, T., Richardson, M., & Wass, V. J. (2009). *We Sell Our Time No More: Workers' Struggles against Lean Production in the British Car Industry*. Pluto Press.

Woods, M., Macklin, R., Dawkins, S., & Martin, A. (2019). Mental Illness, Social Suffering and Structural Antagonism in the Labour Process. *Work, Employment & Society, 33*(6), 948–965.

Disaster Disparities and Differential Recovery in New Orleans

Christina Finch, Christopher Emrich, and Susan Cutter

1 Introduction

The potential for a hurricane to cause massive damage in New Orleans was well-known among disaster researchers and emergency management practitioners with many suggesting that New Orleans was a disaster waiting to happen, it was only *when*, not *if* (Fischetti, 2001; Laska, 2008). Then in August 2005, the world watched in disbelief at the graphic media images of the failed response—residents perched on rooftops awaiting rescue, people stranded in the Superdome, armed military forces patrolling the streets. These horrific images of the social catastrophe tell a poignant story

C. Finch
Pacific Disaster Center, Kihei, HI, USA
e-mail: cfinch@pdc.org

C. Emrich • S. Cutter (✉)
Hazards & Vulnerability Research Institute, Department of Geography, University of South Carolina, Columbia, SC, USA
e-mail: scutter@sc.edu

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_28

of disparities in American cities, and provide an opportunity to learn from them about the disaster resilience of places and the social systems embedded within them. Hurricane Katrina damaged 71% of New Orleans' occupied housing units, making this event the largest residential disaster in US history and afflicting a diverse socioeconomic spectrum of residents with widespread damage. Four years later, Orleans Parish is still on the road to recovery with 31% fewer residents; 46% fewer children enrolled in public schools; 22% less of its labor force; and 48% fewer state-licensed hospitals operating than it had a month before the August 2005 storm (Brookings Institution, 2009).

Many researchers claimed that the disaster response and recovery operations were all steeped in racial bias (Dyson, 2006; Hartman & Squires, 2006; Potter, 2007). However, there were many historical sociodemographic disparities in New Orleans around the nexus of race, class, gender, and age, which led to the social catastrophe known as Hurricane Katrina. The development of New Orleans has been fraught with past environmental injustices, resulting in the most socially marginalized population living in the highest risk areas (Bates & Swan, 2007; Campanella, 2007; Colten, 2005; Fussell, 2007; Germany, 2007). The pre-Katrina social patterns shaped the way residents made decisions to evacuate, how they coped with the events, and how and at what pace they will recover. The social differentiation in the city of New Orleans resonates in every aspect of emergency management from preparedness measures and the initial response to the storm by local, state, and federal officials, to long-term recovery. Socioeconomic stratification and its distribution in the city continue to influence the long-term recovery and mitigation efforts currently underway.

The purpose of this chapter is to measure how these social disparities affect the geography of recovery in New Orleans. Specifically, we examine how the pre-existing social vulnerability of New Orleans neighborhoods intersects with the level of flood exposure to produce spatial inequalities in disaster impacts. Further, we relate these differential impacts based on flood exposure and social vulnerability to the status of recovery, thereby illustrating the socio-spatial disparities in recovery. Lastly we evaluate the spatial distribution of Road Home grants for repairs/ rebuilding, residential sales, and relocations. Our goal is to provide empirically supported, spatial-temporal evidence for addressing where recovery is progressing and for whom.

2 Social Vulnerability

Social vulnerability is defined as the socioeconomic characteristics that influence a community's ability to prepare, respond, cope, and recover from a hazard event (Cutter et al., 2003; Laska & Morrow, 2006). Social vulnerability is the product of social stratification and inequalities—it is not only a function of the demographics of the population but also of more complex constructs such as health care, social capital, and access to lifelines including emergency response (Cutter & Emrich, 2006). Social vulnerability frames disasters and their impacts within broader social contexts and processes (Tierney, 2006; Wisner et al., 2004).

Hurricane Katrina demonstrated the reality of the pre-existing social vulnerability of New Orleans. There was not one single dimension of vulnerability that led to the social catastrophe of Hurricane Katrina; but rather the interaction of multiple dimensions—race/ethnicity, income, family structure, housing—that ultimately created the differential abilities of residents to prepare for, respond to, and now recover from Hurricane Katrina. As seen in the burgeoning literature, the social burdens of the storm's impact were uneven, largely borne by the African American community, the poor, renters, the unemployed, and the undereducated (Brunsma et al., 2007; Elliott & Pais, 2006; Logan, 2006; Masozera et al., 2007; Myers et al., 2008). Similarly, the ability to return and rebuild is also a function of race, class, and gender differences (Landry et al., 2007; Leong et al., 2007a, 2007b; Long, 2007; Pastor et al., 2006). Understanding how pre-existing social vulnerability contributed to the disparities in impacts and eventually to the differential recovery within the city is critical in determining the resilience of New Orleans and likely obstacles in the rebuilding and repopulation of the city (Myers et al., 2008). The empirical evidence and the experiences of New Orleans provide useful lessons that can be utilized in different contexts and locales in preparing for and responding to future disasters.

Social indicators research has a long tradition within the social sciences and provides the methodological rigor for the development of social vulnerability metrics. Coupled with the richness of field-based studies on the impact of disasters on communities and residents, quantitative measures of social vulnerability are now being developed, field tested, and scrutinized for applicability (Birkmann, 2006; Cutter et al., 2003). One of the most prominent approaches is the place-based Social Vulnerability Index (or SoVI), initially constructed to compare US counties based on their social

vulnerability (Cutter et al., 2003). After an extensive literature review and statistical vetting, Cutter et al. (2003) identified 42 input variables that captured the multiple facets of social vulnerability. A principal components analysis (PCA) grouped the underlying, independent variables representing vulnerability into broader categories or constructs (e.g., socioeconomic status, labor force participation, age composition). These components combined in an equal-weighted additive model, which made no a priori assumptions about the relative importance of each component in producing social vulnerability. The resulting "index score" represents the totality of social vulnerability for the enumeration unit (county). Given that much of the prior research focused on individual dimensions or indicators such as race, class, or gender, the development of a composite measure such as SoVI provides both a multidimensional and comparative view of social vulnerability, especially as it manifests itself differently in places. It is in fact the confluence of all the dimensions, which characterizes the social vulnerability of places and its variability over time and across space.

Since its initial debut, replication of the Social Vulnerability Index (or SoVI) has occurred in a number of different contexts. These include an analysis of historic spatial trends in social vulnerability for US counties (Cutter & Finch, 2008) and an assessment of the Gulf Coast counties affected by Hurricane Katrina (Cutter & Emrich, 2006; Myers et al., 2008). SoVI has been utilized to understand hurricane wind risk in Miami-Dade county (Chang & Lindell, 2005) and is used in multiple state (California, Colorado, and South Carolina) and sub-state level (Chakraborty et al., 2005) hazard vulnerability assessments as part of the mandatory FEMA hazard mitigation planning process. Additionally, the social vulnerability index has been applied outside of the US context in Austria (Fuchs, 2009), Germany (Fekete, 2009), Romania (Armas, 2008), and Mexico (Collins et al., 2009). More significantly for this chapter, the SoVI methodology has been downscaled to the Census tract level for the coastal counties of Mississippi (Cutter et al., 2006)[1] and has been cited and discussed in more than 70 peer-reviewed journal articles, including 13 related to the Katrina disaster and 7 specifically discussing the New Orleans area (Borden et al., 2007; Ciger, 2007; Costanza et al., 2006a, 2006b; Curtis et al., 2007; Cutter et al., 2006; Schmidtlein et al., 2008).

[1] Other applications of the Social Vulnerability Index (SoVI) can be found online: http://webra.cas.sc.edu/hvri/products/soviapplications.aspx

3 The Social Vulnerability Index for New Orleans

Utilizing the original methodology outlined by Cutter et al. (2003), a Social Vulnerability Index for New Orleans (SoVI-NO) was created at a Census tract level using socioeconomic data from the 2000 Census. While we attempted to collect the 42 socioeconomic variables outlined in the original methodology, the change in the level of geography, from county to tract, resulted in some of the variables being unavailable.[2] Further, our focus was on the social fabric and social construction of vulnerability, so we omitted indicators of built environment vulnerability. Sensitivity analyses of the SoVI algorithm suggest that SoVI is robust enough to withstand minor changes in variable composition and scale (Schmidtlein et al., 2008). There were 29 variables collected for the 181 Census tracts within Orleans Parish (Table 1). Following the procedures in the original construction of SoVI, a principal components analysis (PCA) defining the underlying, independent, dominant components of social vulnerability for the study area was undertaken. The PCA explained 74.8% of the variance in the original input dataset with seven components (Table 2) combining to describe social vulnerability in Orleans Parish. While each component influences the overall score for the study area to some extent, three dominant components for Orleans Parish are revealed in the PCA, socioeconomic status (race and class), working women with young families (percentage of children under five, and percentage of the labor force who are female), and the public housing developments (income, renters, and density). After adjusting the seven multi-variate components to reflect their influence on vulnerability (positive loadings increase vulnerability; negative loadings decrease vulnerability), they were combined using an additive model to produce the final index score, SoVI-NO. As in the original methodology, the additive model for SoVI-NO made no a priori assumptions about the importance of any single component instead it assumed equal weighting for all the seven components. Though other assessments have specifically identified the importance of individual components of social vulnerability (income, gender, age, for example), the

[2] Cutter et al. (2003) outlined the original 42 variables. The following variables were not available for this analysis: birth rate, percentage voting in the 2000 elections, county debt/revenue, population change, and measures of the built environment: density of residential property, new building permits per square mile, density of manufacturing, density of commercial development, earnings in all sectors, percentage of land in farms, percentage urban, hospitals per capita, and rural farm population.

Table 1 Variables used for the social vulnerability index for Orleans Parish, census tracts (SoVI-NO)

Variable name	*Description*
QBLACK	% Black population
QINDIAN	% Native American population (American Indian, Eskimo, or Aleut)
QASIAN	% Asian or Pacific Islanders population
QSPANISH	% Hispanic persons
QKIDS	% Population under 5 years old
QPOP65O	% Population 65 years or older
MEDAGE	Median age
QFEMALE	% Female
QCVLUN	% Of the civilian labor force unemployed
PERCAP	Per capita income
PPUNIT	Average number of people per household
QRICH	% Households earning more than $ 100 K in 2000
QPOVTY	% Persons living in poverty
MDHSEVAL	Median dollar value of owner-occupied housing units
QRENTER	% Renter-occupied housing units
MEDRENT	Median gross rent ($) for renter-occupied housing units
PHYSICN	Number of physicians per 100,000 population
QFHH	% Families with female-headed households with no spouse present
QMOHO	% Of housing units that are mobile homes
QED12LES	% Population over 25 years old with less than 12 years of education
HODENT	# Housing units per square mile
QCVLBR	% Civilian labor force participation
QFEMLBR	% Female participation in civilian labor force
QAGRI	% Employed in primary industry (farming, fishing, mining, forestry)
QTRAN	% Employed in transportation, communications, and other public utilities
QSERV	% Employment in service occupations
QNRRES	% Nursing home residents
QSSBEN	% Social security recipients
MIGRA	% International migration

purpose of the SoVI index generally, and the SoVINO specifically is to provide an aggregate measure of *all* the factors that contribute to the social vulnerability of the parish, not just one or two individual variables or components.

The SoVI-NO map displays the distribution of social vulnerability within the city of New Orleans (Fig. 1). The SoVI-NO scores were mapped using standard deviations (Std. Dev.) as the classification algorithm to highlight the extremes (low and high) in social vulnerability within New Orleans: areas of low vulnerability (less than -0.5 Std. Dev. from the mean)

Table 2 Dominant components of the Social Vulnerability Index-New Orleans

Component	*Name*	*% Explained*	*Sample variables loading high on the component*
1	Race and class	25.2%	Black, education less than high school, service occupations, poverty, unemployment
2	Young families	12.4%	Female, females in the labor force, % kids under 5, people per housing unit
3	Public housing developments	9.2	Renters, housing unit density
4	Elderly	9.1	Social security recipients, % population over 65, median age
5	Hispanic immigrants	7.4%	Hispanic, % international migration
6	Special needs	6.3%	Nursing home residents, % manufactured housing (mobile homes)
7	Natural resources employment	5.1	% Asian, % employed in extractive industries (fishing, farming, forestry, mining)
	Total	74.8	

and high vulnerability (greater than 0.5 Std. Dev. from the mean).[3] As expected, the majority of Census tracts (56.3%) fall within the medium range of vulnerability, with the remaining tracts divided between the low (19.9%) and high (23.8%) vulnerability categories.

The neighborhoods (Census tracts) classified as "Low" (white) have lower social vulnerability scores in comparison to the rest of the city. This recognizes that certain underlying socioeconomic and/or demographic characteristics of these areas would enable a community to better prepare for, cope with, and recover from a hazard event. The Touro neighborhood (http://www.gnocdc.org/orleans/2/67/snapshot.html) is a good example of an area with low-social vulnerability, prior to Hurricane Katrina. Touro, a relatively small residential neighborhood (3242 residents in 2000), contains historic plantations (Macarty and Louis Bouligny) and medical facilities (Touro infirmary and St. Charles Specialty Hospital). Lawrence Park, located within Touro, is an active recreation area with

[3] The neighborhood delineations were developed by the Greater New Orleans Community Data Center (http://www.gnocdc.org/def/neighborhood.html). The neighborhood boundaries were used for visualization purposes only; all analyses were performed on the Census tract level.

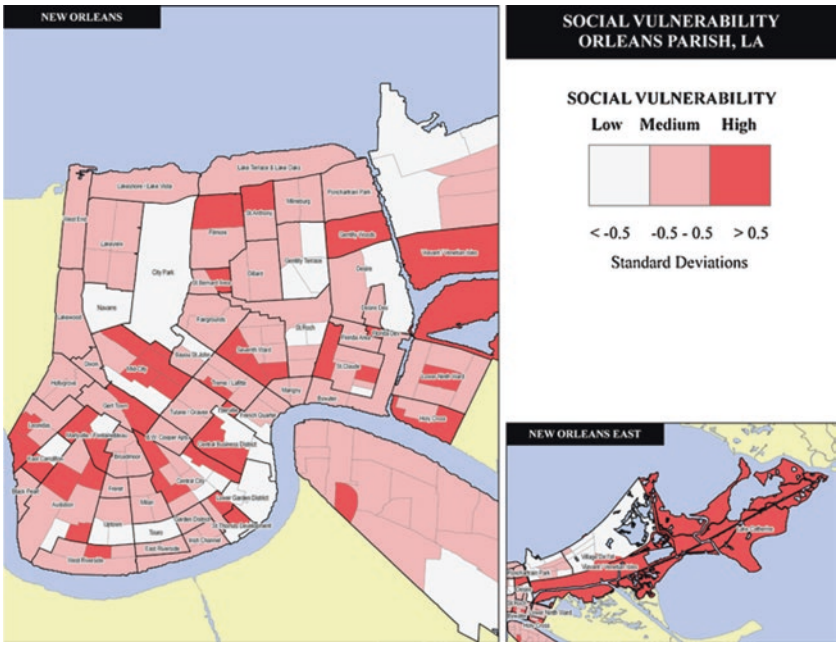

Fig. 1 The Social Vulnerability Index for New Orleans (SoVI-NO). (Color figure online)

playgrounds and multipurpose fields. There is some racial diversity in Touro, but the majority of residents (75%) are white. As measures of wealth, 15.5% of the residents live below the poverty level, while more than 10% of the residents can be considered "wealthy" (households earning more than $100,000) with an average household income of $48,145. More than 85% of the residents have a high school diploma, and the unemployment rate was around 1%. The Touro neighborhood as well as others in this low-social vulnerability category, would be more likely to have the available resources to recover from a hazard event within shorter time periods.

At the other end of the spectrum, the areas represented in the "High" category (red) have the highest levels of social vulnerability. In these communities, residents would struggle to prepare for, cope with, and recover from a hazard event, largely due to their pre-event socioeconomic circumstances (see previous section). Many of the high-social vulnerability areas

correspond with locations of public housing developments managed by the Housing Authority of New Orleans, such as the C.J. Peete (Central City); B.W. Cooper Apartments; St. Thomas Development; St. Bernard; Desire Development; Florida Development; Lafitte (Treme/Lafitte); and the Iberville (Iberville).[4] In addition, residents of public housing developments, like renters, are unable to control the rehabilitation of their housing status after a disaster. Furthermore, residents of these areas are more likely to be minority populations struggling with lower incomes, limited education, unemployment, and a lack of access to political power. An example of a neighborhood in the high category of social vulnerability is Holy Cross. Holy Cross (population of 5507 in 2000) is a medium-sized neighborhood in New Orleans similar to Touro, yet the socioeconomic characteristics vary drastically. Holy Cross is 87.5% African American, with 7.1% of its residents unemployed and 29.5% of the residents living below the poverty level. More than 25% of the households are female-headed (with no spouse present), compared to 2% for Touro. Only 3.3% of the households qualify as wealthy (more than $100,000 in income), and the average household income is $32,202. In addition, 36.4% of Holy Cross residents have less than a high school diploma.

Due to the complex and aggregate nature of the social vulnerability metrics in the SoVI-NO, the driving forces contributing to high-social vulnerability vary from place-to-place. In contrast to Holy Cross, for example, the Filmore neighborhood is middle to upper class, which would lead one to expect a low-average level of social vulnerability. However, the northern tracts in Filmore are classified as High-social vulnerability. In this area, the driver is age, specifically elderly populations. Approximately 25% of the residents are over 65 years old and 18% of the residents are recipients of social security benefits. Elderly populations often have mobility and health issues which require extra resources during emergency situations—causing this population to be disproportionately more vulnerable in comparison to other age groups (Ngo, 2001; Cutter et al., 2000; Hewitt, 1997; O'Brien & Mileti, 1992). The northern census tract in the Filmore neighborhood contains a nursing home facility whose residents would require special assistance for disaster preparation and a likely evacuation. Such residents have many special needs (medications and medical

[4] The Housing Authority of New Orleans was in the process of destroying, renovating, and relocating some of the housing projects pre-Katrina; however, this is not reflected in the SoVI-NO because it is based on data collected from the 2000 Census.

equipment), and may require assistance with transportation out of harm's way (Morrow, 1999; Tobin & Ollenburger, 1993).

Other neighborhoods in the greater New Orleans area illustrate the complexity of the Social Vulnerability Index. Despite its affluence, East Carrolton contains pockets of vulnerable populations in selected census tracts due to the percentage (34%) of renters and female labor force participation (50%). Finally, the Audubon neighborhood, composed of five census tracts, has substantial variability in population characteristics, including persons living in poverty (between 7 and 25%) and higher housing densities (from just over 1000 per square mile to more than 4500), all that contribute to the different levels of vulnerability within the neighborhood when using tract level census data.

The relative levels of social vulnerability within Orleans Parish provide a broad overview based on aggregate measures of which areas contain more socially vulnerable populations than others. Still, SoVI-NO is a place-based tract level (neighborhood) assessment of pre-event conditions that are known to influence and exacerbate the adverse impacts from disasters. SoVI-NO provides insight into the complex nature of underlying socioeconomic and demographic conditions as a means to understand the geographical manifestation of differences in abilities to prepare for, respond to, and recover from disaster events. As such, SoVI-NO does not reflect the capacity of specific individuals or families to return and recover from Katrina, but does provide insight into broader spatial trends in recovery from the storm at the neighborhood scale.

4 Hurricane Katrina's Flood Waters

Floodwaters filled the city of New Orleans as Hurricane Katrina breached the levees. On September 11, 2005, a day of near-maximum flooding in New Orleans, the Federal Emergency Management Agency (and its federal counterparts) used remote sensing imagery to assess the extent of the flooding and the depth of the floodwater (FEMA, 2005). These data were the primary resource for flooding data post-Hurricane Katrina; they were also used by the Greater New Orleans Community Data Center (GNOCDC) for their flood extent map.[5] The flood inundation data

[5] One problem to note, since the imagery measures standing water, areas such as Lake Catherine, which had significant damage due to storm surge, but no standing water, are not represented in the final dataset.

included the mean depth of the floodwater for each Census block. To facilitate comparisons between the SoVI-NO and the magnitude of flooding we computed the mean flood depth for each Census tract by averaging the mean flood depth from every Census block within each tract. Using the mean flood depth rather than the maximum flood depth for each census block provides a conservative flood depth value for each census tract that can then be compared to the social vulnerability scores and recovery rates.

The map (Fig. 2) classifies the flood inundation at a Census tract level based on the average flood depth: None (0), Low (<2 ft), Medium (2–4 ft), and High (>4 ft). These flood depth classifications approximate levels of damage: Low or minor damage, Medium or serious damage, and High or severe damage (McCarthy et al., 2006). The flood depth and damage equivalent are based on US Army Corps of Engineers studies, which assess

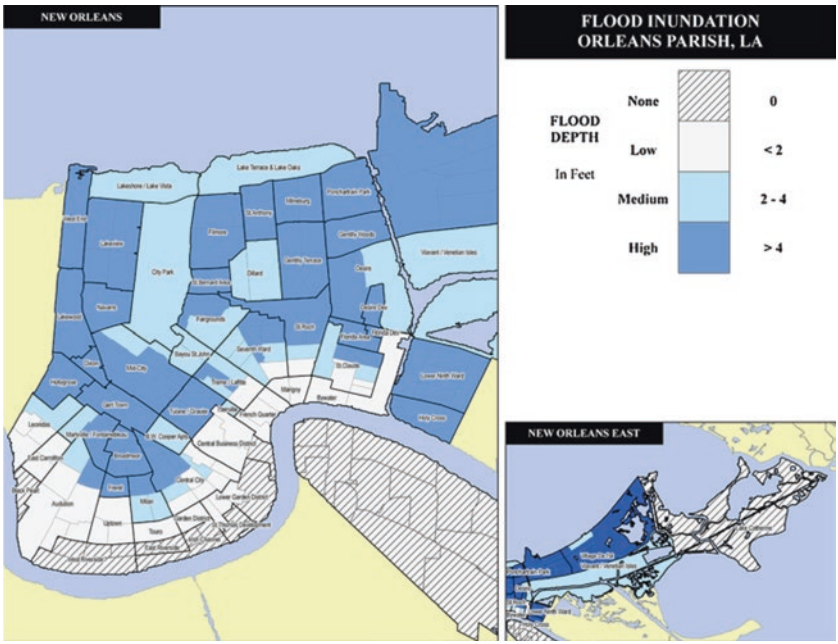

Fig. 2 Hurricane Katrina Flood Inundation for Orleans Parish, LA by Census Tract. Low is less than 2 ft., medium is 2–4 ft., and high is >4 ft. Cross-hatched areas experienced no flooding. (Color figure online)

typical structural damage produced by standing water, which states that water in excess of 3 ft. typically produces extensive structural damage (McCarthy et al., 2006). We used 3 ft. as the mid-range for the moderate damage category suggesting that anything greater than 4 ft. (the next integer) would assuredly have severe-catastrophic structural damage. Thus, in the absence of damage estimates for each individual structure, flood depth is a surrogate used to estimate the potential damage. Tracts were classified as follows: 17% of the tracts had no flooding; 22% of the tracts had less than 2 ft. of flooding (Low or minor damage); 17% had between 2 and 4 ft. of flooding (Medium or serious damage); and 44% of the tracts had more than 4 ft. of flooding (High or severe damage).

5 The Makings of a Disaster and Uneven Recovery

Recovery from a disaster is a function of the magnitude of the disaster's impact on places, the pre-existing vulnerability and coping capacity of the affected population, and access and availability of recovery assistance. The interaction of the physical and social processes provides the basis for examining the differential burdens of the Katrina disaster. Figure 2 clearly illustrates that the majority of New Orleans neighborhoods (83%) experienced some flood damage, irrespective of elevation or socioeconomic characteristics. Yet, when the flood levels are combined with the social vulnerability, a very different geography of impact emerges (Fig. 3). Areas with both low inundation and low-social vulnerability (light gray) include the Lower Garden District, Touro, and Uptown. We would expect these neighborhoods to be among the first sections of the city to recover given the minimal amount of flooding and higher neighborhood ability to prepare for, respond to, cope with, and recover from the disaster based on the relatively low levels of social vulnerability. On the other hand, areas in dark blue and red to black had extensive flooding as well as some of the highest levels of social vulnerability. In these neighborhoods (Mid-City, the Lower Ninth, Gert Town, and Holy Cross) extreme physical damage coupled with socioeconomic characteristics indicating high-social vulnerability, reflects a pattern that would most likely hinder response and recovery efforts. In examining this 4 by 3 matrix, we can view the spatial distribution of the overall impact of Katrina on neighborhoods, but we can also use this map to determine what factors (physical or social) are most likely to drive the recovery process. For example, the timeline of recovery in neighborhoods with medium-high levels of flooding and low-medium vulnerability (Gentilly Terrace, Lakeview), is more likely to be influenced

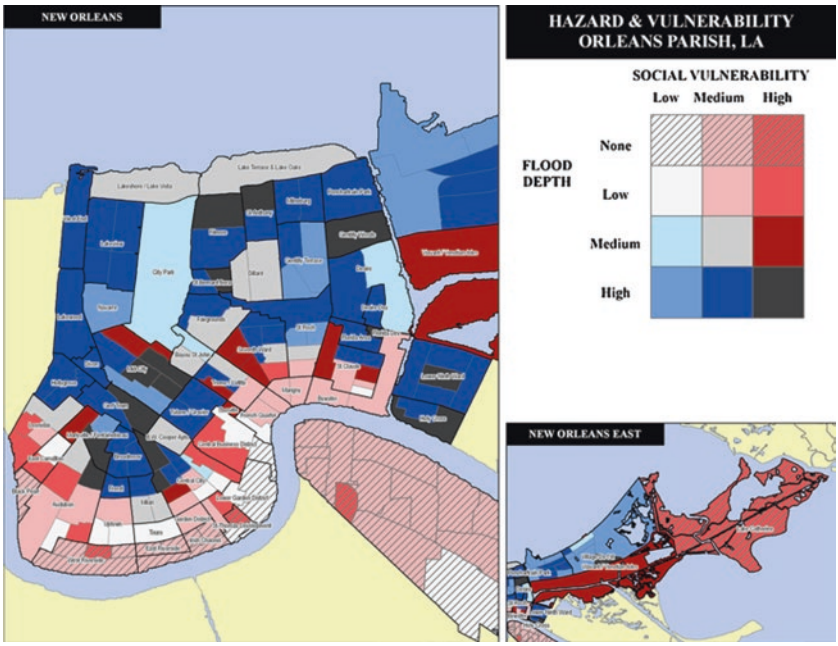

Fig. 3 Intersection of social vulnerability and flood inundation

by the depth of flooding rather than the social vulnerability. Though these areas had significant flooding, their socioeconomic characteristics (lower social vulnerability) provide a greater capacity to recover faster than elsewhere in the city. Areas with medium levels of flooding, but high-social vulnerability (Viavant/Venetian Isles, Treme/Lafitte, Iberville) portend a potentially longer recovery period, largely due to the social vulnerability.

The geographic variability in impacts produces disparities in the recovery. Fieldwork in the aftermath of Hurricane Katrina is documenting areas experiencing different roads to recovery, despite equal levels of devastation based on flood inundation. Researchers commented on the role of pre-existing social conditions in the ability of some economic classes to respond and cope with the aftermath of Hurricane Katrina (Barnshaw & Trainor, 2007; Elder et al., 2007; Elliott & Pais, 2006; Leong et al., 2007a, 2007b; Long, 2007; Masozera et al., 2007; Miller & Rivera, 2007; Potter, 2007; Trujillo-Pagan, 2007). However, these studies do not provide the geographic detail necessary to understand the spatial differences in recovery throughout the entire Parish in relation to the level of flood

damages. Similarly, detailed comparisons of neighborhoods suggest an uneven pattern of recovery as well. For example, in their comparison of Lakeview and the Lower Ninth Ward, Elliott et al. (2007) found Lakeview residents had electricity and phone service, were more likely to return to the pre-Katrina status, and had more neighborhood organization than the Lower Ninth Ward inhabitants. As seen in Fig. 3, Lakeview and the Lower Ninth Ward experienced high levels of flooding; however, they had differing levels of social vulnerability, which could account for some of the differential recovery. The comparison of Central City and Treme highlighted differences in the recovery population by age, household size, race, income, and homeownership (Mock et al., 2007), or some of the individual variables that produce social vulnerability. Finally, a comparison between the Upper and Lower Ninth Wards illustrates the impediments in recovery based on housing structure and level of damage (Green et al., 2007). Despite these laudable research efforts, most are localized case studies of one or two communities. A broader view of the spatial inequalities in recovery is still lacking. Such an analysis is important because it provides a unique way of analyzing disparities across a landscape in comparison to other areas based on pre-event characteristics and impacts from the hazard. While utilizing spatial analytics to understand trends in populations and hazard impacts is not (by itself) new or novel, the inclusion of space as a required element of this analysis does provide a new research angle in the study of long-term recovery from disaster events. Using space as a factor in disaster rebuilding and recovery research provides a completely new and unexplored realm which could be transitioned into other fields of research in the human-environment arena.

6 Tracking Repopulation and Recovery

One of the major constraints on tracking the recovery of New Orleans is the lack of reliable estimates of the number of residents in the city, who is returning, who is not, and where are they locating. Most recent US Census estimates are for July 2007 and place the Orleans Parish population at 288,113 or 63% of pre-Katrina levels (July 2005) (Greater New Orleans Community Data Center, 2008).[6] These numbers are in flux with people

[6]The July 2007 population estimates for Orleans Parish reflect Census Bureau revisions based on successful challenges by the parish government. The July 2007 population estimate was revised upward from 239,124 to 288,133. The July 2005 and July 2006 population estimates for Orleans Parish also will be revised upward, although those revisions have not been released by the Census Bureau at this time.

leaving temporarily or indefinitely, displaced residents returning to the city, and newcomers arriving to rebuild the city.

In order to validate how well the social vulnerability and flood damage model predicts recovery in New Orleans, we need a consistent measure of population at the neighborhood level. Unfortunately, there are many shortcomings in the existing data sets, especially in the scale of analysis and the timeliness of population estimates. For example, the US Census Bureau estimates demographic data for counties on an annual basis using the American Community Survey (ACS), but this is at the county or parish level only (Frey et al., 2007). Single purpose snapshots are available such as those done immediately after Hurricane Katrina for the Louisiana Public Health Institute, but again this was only at the parish level (Louisiana Public Health Institute, 2006). The most consistent sub-parish source of data is from the Brookings Institution and the Greater New Orleans Community Data Center (GNOCDC). Their Katrina Index has been tracking the recovery process through many indicators including population, housing, economy, education, and infrastructure. Initially published monthly online (www.gnocdc.org), the Katrina Index has transitioned to the New Orleans Index and will be compiled quarterly and focus on the City of New Orleans and its surrounding parishes (Brookings Institution, 2009). The indicators by Brookings are available from August 2005 to December 2008 with indicators summarized for the state, metropolitan area, or parish. Some sub-parish level data for the 13 planning districts track the repopulation for these areas, but there is no finer grained data at neighborhood level.

In response to the demand for more detailed and timely population estimates, the GNOCDC evaluated the potential of using statistics from the US Postal Service (USPS) as a measure of population recovery (Plyer & Bonaguro, 2006). The USPS compiles a delivery statistics product monthly detailing the number of residences actively receiving mail for each postal carrier route. As noted by Plyer and Bonaguro, there are limitations with these data: carrier route boundaries do not align with US Census boundaries; carrier route boundaries change over time; and there has been no validation of how well residential deliveries of mail correlate with repopulation. To address some of these issues, Valassis Direct Mail Inc. works cooperatively with the USPS to review and improve the quality of the database and then provide updated data to the GNOCDC (Ortiz & Plyer, 2008). The GNOCDC continues to utilize and review the Valassis database, therefore, in the absence of other data on repopulation, the

active residential deliveries are a good proxy for occupied households and can help track the progression of recovery based on repopulation of the New Orleans neighborhoods.

For this analysis, we procured the USPS active residential deliveries data from Valassis Lists through a data download portal at the GNOCDC for two time periods: June 2005 and June 2008 coinciding with two months before Katrina and three years afterward.[7] We aggregated the residences actively receiving mail to the Census tract level for the entire study area, to enable comparisons with depth of flooding and social vulnerability (SoVI-NO). At present, interim data (e.g., June 2006, June 2007) are not available, thus we base our estimate of recovery on a three-year cumulative post-event return rate using residential postal deliveries from 2005 as our base year. The return rates are simply derived from the total number of postal addresses in the census tract receiving mail before and after Hurricane Katrina. They do not reflect who is receiving the mail or whether the same residents were there pre-storm, or if there is a different resident in the structure post-event. The return rates should be viewed as an aggregate measure for the tract in same way that we aggregate population characteristics.

7 The Places and Faces of Recovery

Flooding from Hurricane Katrina and the failure of the levees inundated 83% of the Census tracts in New Orleans Parish. The inundation area included tracts with all three designations of social vulnerability—High, Medium, and Low. We found that the level of flooding was not significantly correlated with social vulnerability (Pearson's $r = 0.03$, significance = 0.70), demonstrating that Hurricane Katrina affected all areas, regardless of socioeconomic composition or social vulnerability. There is, however, a spatial and statistical difference in how those neighborhoods are recovering.

In July 2005, the USPS reported 203,181 active residential deliveries in Orleans Parish, LA. Three years after Hurricane Katrina (June 2008), there were 146,158 residential deliveries for an estimated return rate of 71.9%. We expected that flood level determined the amount of destruction and thus influences the likelihood of return. There was a strong

[7] Data driven by Valassis Lists. From a compilation by the Greater New Orleans Community Data Center, http://www.gnocdc.org, December 2008.

negative correlation between the percentage of households returning and the mean depth of flooding, suggesting that tracts with higher flooding had lower percentages of returned households (Pearson's $r = -0.69$, significance = 0.001) (Table 3). After three years, for example, neighborhoods with no flooding had a return rate of 96% (Table 4). The Low flood depth category (less than 2 ft. inundation), also shows a strong recovery as well with 96.2% of households receiving mail in June 2008. In the Medium flood depth areas (2–4 ft. inundation), the June 2008 return rate was 77.7%. Finally, in those Census tracts with flooding greater than 4 ft., only 52.4% of the households received mail deliveries three years after Hurricane Katrina. As expected, the increased depth of flooding (and thus increased damage) significantly decreases the likelihood of recovery (or return). Therefore, the geographic pattern of the percentage of returning households three years after Hurricane Katrina (Fig. 4) is very similar to the flood inundation map (Fig. 2).

When examining the percentage of returned households and pre-existing social vulnerability, some interesting patterns emerge. Three years after Hurricane Katrina, 16 out of the 50 neighborhoods in New Orleans had less than 50% of the households receiving mail (Brookings Institution, 2009). In our analysis, 15 Census tracts, such as those in the Lower Ninth Ward, St. Bernard, and B. W. Cooper Apartments, have less than one-third of the pre-Katrina residents receiving mail (Fig. 4). The SoVI-NO is significantly correlated with the percentage of returned households (Pearson's $r = -0.25$, significance = 0.001). This negative relationship suggests that as the level of social vulnerability increases the percentage of households returned decreases, in other words return rates in neighborhoods with the highest social vulnerability are lower than return rates in neighborhoods with low to moderate levels of social vulnerability.

Table 3 Returnees, flood levels, and social vulnerability

	Social vulnerability	*Mean flood height*	*% Households returned*
Social vulnerability	–	0.03	-0.25**
Mean flood height	0.03	–	-0.69**
% Households returned	-0.25**	-0.69**	–

Pearson's *r* was used as the correlation statistic

** Significant at $p = 0.001$

Table 4 Average percent returned based on social vulnerability and level of flooding

Percent returned	*Social vulnerability*			
	Standard deviations			
	Low < -0.5	*Medium -0.5 to 0.5*	*High > 0.5*	*All levels*
Average flood depth				
In feet				
None				
June 2005	5536	26,301	2087	33,924
June 2008	5373	25,215	1992	32,580
% returned	97.06***	95.87***	95.44*	96.03***
# of tracts	4	23	4	31
Example	Eastern lower Garden District	East Riverside, Irish Channel	Lake Catherine	
Low <2				
June 2005	8736	20,366	8795	37,897
June 2008	9090	18,973	8391	36,454
% returned	104.05***	93.16***	95.41***	96.19***
# of tracts	11	18	11	40
Example	Touro, lower uptown	Garden District, Marigny, Bywater	Central BD, South Leonidas	
Medium 2–4				
June 2005	6108	16,956	9541	32,605
June 2008	4808	13,361	7179	25,348
% returned	78.72*	78.80***	75.24***	77.74***
# of tracts	5	14	11	30
Example	City Park, north bayou St. John	Dillard, B.W. Cooper Apts	Viavant/venetian isles, west 7th Ward	
High >4				
June 2005	24,609	56,213	17,933	98,755
June 2008	14,637	27,715	9424	51,776
% returned	59.48***	49.30***	52.55***	52.43***
# of tracts	16	47	17	80
Example	Navarre, village de L'est	Ponchartrain Park, Lakeview, Milneburg	Gentilly woods, south holy cross	

(*continued*)

Table 4 (continued)

Percent returned	*Social vulnerability*			
	Standard deviations			
	Low < -0.5	*Medium -0.5 to 0.5*	*High > 0.5*	*All levels*
All levels				
June 2005	44,989	119,836	38,356	203,181
June 2008	33,908	82,264	26,986	146,158
% returned	75.37***	71.15***	70.36***	71.93***
# of tracts	36	102	43	181

Significance of difference of proportions test (2-tailed *t*-test): * $s = 0.05$, ** $s = 0.01$, *** $s = 0.001$

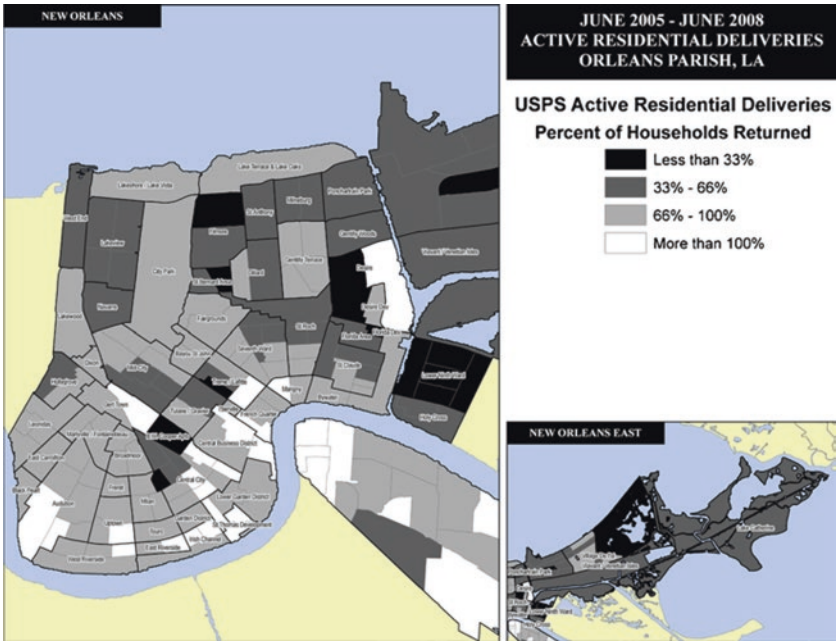

Fig. 4 Percentage of active residential deliveries in New Orleans, 3 years after Hurricane Katrina

The pattern of recovery becomes even more significant when evaluating the social vulnerability in conjunction with the level of flood depth. Communities with little or no flooding and low levels of social vulnerability averaged 93–100% of the households returned. Neighborhoods with medium levels of flooding and high social vulnerability (Vivant/Venetian Isles) had return rates of 75%. Finally, areas with the most significant flooding (more than four feet) and high-social vulnerability (e.g., tracts within Gentilly Woods or Holy Cross) had approximately 50% of the households returning after three years (Table 4).

Perhaps the most telling statistic is that the slowest recovery is among the Census tracts in the Medium range of social vulnerability and the highest flood levels, where only 49.3% of households have returned (Table 4). Located in areas such as Ponchartrain Park, these Census tracts contained significant working-class residential areas with single-family owner-occupied homes. While media, celebrity support, and federal programs focused on communities such as the Lower Ninth Ward, assistance for the working-class households has been slow in materializing, and thus their recovery period may be prolonged. In many respects, they are the in-between population—not poor enough to qualify for outright assistance, yet too poor to recover based on their own assets, including insurance. We speculate that these working-class households, mainly African American, are the "forgotten casualties" of Katrina, residents who are shouldering much of the disaster's impact in the long term.

8 Governmental Support for Recovery

To test the assertion of the forgotten casualties, we examined the governmental support for recovery using two indicators from the congressionally funded State of Louisiana Road Home Program, including: Data from (Option 1) of the Road Home Program, which provides financial assistance to homeowners (up to $150,000) who chose to stay and rebuild in the impact area; and data from (Option 2) or (Option 3) of the Road Home Program which provides (respectively) either financial incentives to purchase another home in Louisiana or fiscal incentives to sell your home and choose not to remain a homeowner in the state. The number of homeowners who received Road Home grants to rebuild their homes was obtained at the Census block level for those pre-Katrina homeowners who closed on their Road Home application and chose Option 1 as of September 18, 2008 (http://www.gnocdc.org/datafile_request/). The

number of damaged or delinquent properties that were purchased by the state of Louisiana by Census block for those pre-Katrina homeowners who closed on their Road Home application as of September 18, 2008 and chose Option 2 or 3, was obtained from the same source. A count of these indicators was aggregated to the Census tract for comparative purposes.

A total of 37,839 properties were supported under the Road Home Option 1 program (repair of damaged houses), while the state buyout (Options 2&3) numbered 4212. As one might expect, the overwhelming majority of the state buyouts (91%) were located in the high flood zone (>4 ft. of flooding). In the case of the Road Home repair, 71.8% of the applicants were in the high flood category. Interestingly, 8.5% of the properties in the Option 1 (Road Home repair) program were located in areas with no flooding. Generally speaking, the distribution of governmental support for recovery appears to track well with the highest levels of flooding and damage, where the mean flood level is positively and strongly correlated with Road Home Option 1 recipients (Pearson's $r = 0.51$, significance = 0.000) and with state buyouts of residential property (Pearson's $r = 0.68$, significance = 0.000).

However, there is quite a different story with respect to the levels of social vulnerability and the distribution of governmental support for recovery (Table 5). The majority of residences for both Option 1 (56.5%) and Options 2 and 3 (60.8%) are located in the Medium category of social vulnerability, which does not support our hypothesis on the forgotten casualties, at least initially. For the most vulnerable, governmental support for repair and buyout was limited (15.3 and 17.0%, respectively). This is easily explained as many of the most vulnerable residents are not homeowners, a targeted population of the particular stimulus program we used (Road Home). The correlation between social vulnerability scores (SoVI) and Road Home Option 1 is negative but weak (Pearson's $r = -0.123$, not

Table 5 Governmental support for homeowner recovery, return rates, flooding, and social vulnerability

Support program	*Mean flood level*	*SoVI score*	*Percent returned*
Road home repair grant (option 1)	0.509***	-0.123	-0.374***
Road home buyout (options 2 and 3)	0.684***	-0.083	-0.556***
All support	0.553***	-0.122	-0.414***

*** Significant at $p > 0.001$

significant) as is the correlation between SoVI and the state buyout program (Pearson's $r = -0.083$, not significant), indicating that with increased levels of social vulnerability, there was less governmental support as measured by the Road Home effort (Table 5).

When examining the pattern geographically, a different interpretation is clear. The governmental support variables were combined, standardized (e.g., divided by the total number of owner-occupied housing units), and then classified into categories (using standard deviations) for mapping purposes. A bivariate map illustrating the levels of social vulnerability and the level of combined governmental support (Road Home and state buyout) shows some of the most vulnerable communities are also receiving the most governmental support, such as Holy Cross, Lower Ninth Ward, St. Bernard Area, and Gert Town (Fig. 5). Similarly, areas of low-social vulnerability received little or no governmental support for recovery, a pattern we expected. However, there are some neighborhoods that

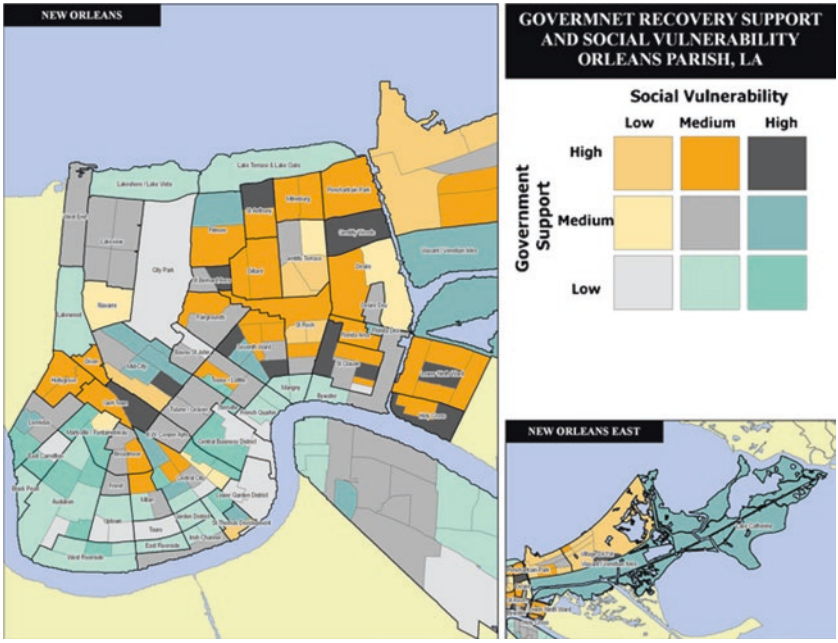

Fig. 5 Government recovery support (Road Home program) and social vulnerability

received high levels of government recovery support but do not necessarily have high levels of vulnerability. Places like Village de L'est and St. Roch received high levels of government support but are among the lowest in terms of social vulnerability, but did have extensive flooding. Conversely, areas such as Mid-City, Iberville, or Viavant/Venetian Isles received relatively low levels of government support despite having higher than average levels of social vulnerability, with moderate levels of flooding.

So how effective is the governmental assistance in stimulating recovery based on the Road Home program? There is a negative correlation between governmental support and return rates/recovery for both the purchase program by the state of Louisiana (Pearson's $r = -0.556$, significance = 0.000) and Road Home repair (Pearson's $r = -0.374$, significance = 0.000). This negative relationship is consistent for all levels of vulnerability and for all levels of flooding. In other words, the number of Road Home grants does not appear to be increasing the number of returning households at the neighborhood level (as measured by postal deliveries) and may in fact provide a disincentive for residents to return to the old residence. But is this true everywhere in the parish?

To see whether there was any spatial variability in the relationship between governmental intervention and postal recipients, we categorized the governmental support and the postal recipients into low, medium, and high categories using standard deviations (Fig. 6). It is very clear from the map that three areas, Lower Ninth Ward, Florida Area, and New Orleans East (Village de L'est), are not recovering with return population as quickly as other sections of the city despite high levels of governmental intervention. On the other hand, two areas, one with significant flooding (Gert Town), and one without (St. Thomas Development) both received significant governmental resources and are recovering at rates faster than others (e.g., 93% for Gert Town and 162% for St. Thomas Development) compared to 11.2% for the Lower Ninth Ward, for example. Some of the difference is due to the buyout and subsequent abandonment of residential areas within the Lower Ninth Ward. Instead of building in the perceived area of high flood potential adjacent to the Industrial Canal, homeowners are seeking homes in other neighborhoods within the city, or not returning at all. At the same time, other communities such as St. Thomas Development are receiving a large influx of residents because of the closure of other public housing developments elsewhere in the city.

There are certain areas throughout the city that show substantial recovery although they have received relatively few Road Home grants or state

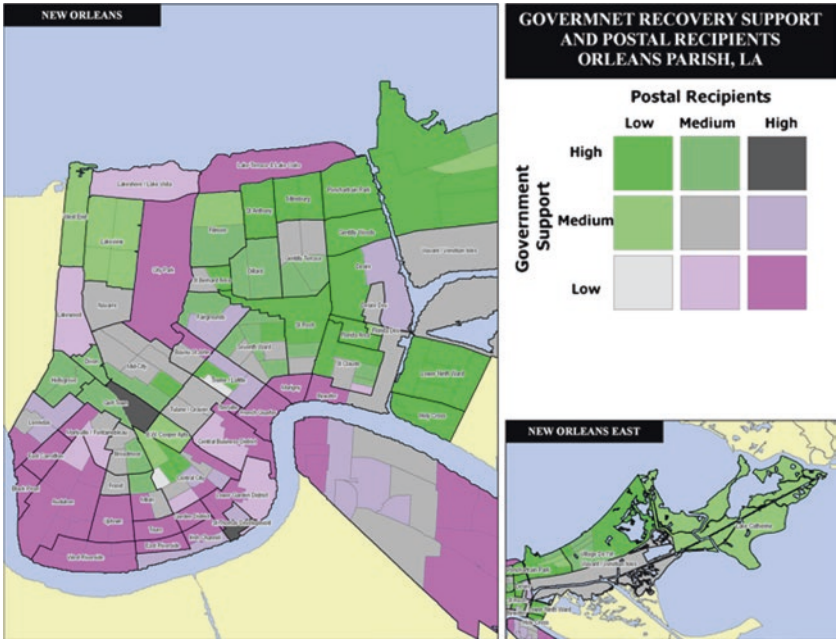

Fig. 6 Relationship between governmental recovery support (Road Home program) and returnees based on postal deliveries

buyouts. Most of the areas in the city bordering the Mississippi River (Marigny, Garden District) or along Lake Ponchartrain have higher levels of recovery based on postal delivery and have received lower than average levels of government support from the Road Home program. This is confirmed by a simple regression model with percentage returning as the dependent variable and flood height, SoVI score, Road Home sales and relocations, and Road Home repairs as the independent variables. The overall model is significant with mean flood height (β = -0.54), SoVI score (β = -0.25), and Road Home sales (β = -0.23) as the most important predictors (s = 0.01). The Road Home repair grants (β = 0.03) were an insignificant contributor to the prediction of return rates. Areas such as these indicate the desire of residents to return to normalcy—with or without the help of government recovery assistance.

9 Discussion and Conclusions

There are clear disparities in the patterns of Hurricane Katrina's impact on New Orleans and its residents. Communities of all income levels and races were flooded illustrating that the physical manifestation of the disaster had few race or class distinctions. On the other hand, the social vulnerability of communities influenced the pre-impact response (who evacuated, and who remained). The social vulnerability also relates to who is returning and how quickly neighborhoods are recovering, but it is less of a statistical predictor of overall recovery than flood height for the entire study area. There are subtle differences between Census tracts when comparing social vulnerability and the percentage change in returned households from June 2005 to June 2008. For example, there is a weak but negative correlation between return rates and social vulnerability, suggesting that communities with higher social vulnerability scores have slower return rates (as measured by postal deliveries) than communities with lower social vulnerability scores. Differences in recovery based on the level of flood inundation, however, are more pronounced and statistically significant (Pearson's $r = -0.69$, significance = 0.001) with areas receiving more than four feet of water recovering slower than those with less flooding. Within the areas with the deepest flooding (>4 ft), it is the medium social vulnerability category, containing neighborhoods such as Lakeview or Milneburg, which have the lowest overall recovery rate.

The combination of high-social vulnerability and high-flood inundation has a distinct spatial pattern that graphically illustrates the social burdens of the Katrina catastrophe, and highlights those areas with lagging recovery. While every neighborhood is showing signs of recovery, those places where the combined impact of the storm in terms of flood levels and the pre-existing levels of social vulnerability coalesce to create a synergistic effect. While external resources from charities have facilitated the recovery of some of the most socially vulnerable neighborhoods such as the Lower Ninth Ward, it is the in-between neighborhoods, not poor enough to qualify outright for assistance, but too poor to recover using their own resources that are lagging. In many cases, it is the neighborhoods with working professional class residents (school teachers, health professionals) in New Orleans that are slower in their recovery. This does not bode well for the parish, where the inequalities between wealthy and poor are becoming more pronounced.

What the analysis does not provide, however, is the underlying causes that are contributing to such disparities in recovery based on the outcome measure we used. We know, for example, that of the roughly 87,600 owner-occupied housing units (pre-Katrina), about 41% are rebuilding based on Road Home approved applications. Another 4.5% have sold their damaged homes to the state of Louisiana, while the remaining 54.3% of owner-occupied homes have no resolution as of March 2008 (www.gnocdc.org). Based on our statistical and spatial analyses, we conclude that the Road Home program has not stimulated recovery and in many communities (e.g., Lower Ninth Ward and New Orleans East) has provided an additional challenge to those wishing to return. While the concept of the Road Home program is soundly based on the principle of helping impacted communities to recover, there are many caveats within its provisions that restrict large segments of the New Orleans community from taking advantage of the program. For example, the buyout provision does not provide a means of rehousing individuals once they have decided to sell. Even more significantly, the Road Home program completely ignores the role of renter-occupied units, the majority of housing tenure options before Katrina. The forced migration of many low-income residents and their inability to return (regardless of desire) may contribute to the low repopulation and partially explain the pattern in the Florida and St. Bernard areas as well as sections of the Lower Ninth Ward. The difficulties in acquiring state and federal assistance for rebuilding, the peculiar requirements of the Road Home program (homeowners only, not renters), and the political dysfunction of the city itself all are contributing to the difficulties in adequately measuring the social and spatial inequalities of disaster recovery.

We have empirically verified using our outcome measure of mail deliveries that the recovery in New Orleans is progressing faster in the less vulnerable and less flooded Census tracts. Those socially vulnerable tracts and those with the greatest flooding are lagging behind in the repopulation of those areas, a finding we expected. When examining the Census tracts with the highest levels of flooding, there are differences in the level of return based on social vulnerability. Tracts at both ends of the social vulnerability index (high and low) are recovering faster than tracts within the middle range of social vulnerability, suggesting that individual resources (for the low socially vulnerability) and governmental/private sector support for the most vulnerable are facilitating a faster recovery in these areas. Yet, governmental support is inversely related to returnees regardless of social vulnerability or flood levels. It is quite clear than many

of the pre-Katrina residents are not returning to their original homes, many of which were purchased by the state. The residential pattern of New Orleans is changing and the new residential landscape portends a future where the geography of recovery is based on who can afford it, not necessarily who was most affected by the disaster based on flooding levels or social vulnerability.

Acknowledgments A preliminary version of this chapter was presented at the Disaster and Migration Conference at Tulane University, April 12–14, 2007. We would like to thank the participants at the workshop for their comments on the research. We also would like to thank Melanie Gall for her critical review of our analysis. Any errors, however, remain ours.

References

Armas, I. (2008). Social Vulnerability and Seismic Risk Perception. Case Study: The Historic Center of the Bucharest Municipality/Romania. *Natural Hazards, 47*(3), 397–410.

Barnshaw, J., & Trainor, J. (2007). Race, Class, and Capital Amidst the Hurricane Katrina Diaspora. In D. L. Brunsma, D. Overfelt, & J. S. Picou (Eds.), *The Sociology of Katrina: Perspectives on a Modern Catastrophe* (pp. 91–105). Rowman and Littlefield.

Bates, K. A., & Swan, R. S. (Eds.). (2007). *Through the Eye of Katrina: Social Justice in the United States.* Carolina Academic Press.

Birkmann, J. (Ed.). (2006). *Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies.* United Nations University Press.

Borden, K. A., Schmidtlein, M. C., Emrich, C. T., Piergorsch, W. W., & Cutter, S. L. (2007). Vulnerability of US cities to Environmental Hazards. *Journal of Homeland Security and Emergency Management, 4*(2), Article #5.

Brookings Institution. (2009). *The New Orleans Index: Tracking Recovery of New Orleans and the Metro Area.* Retrieved August, 29, 2009, from http://www.gnocdc.org/NOLAIndex/NOLAIndex.pdf

Brunsma, D. L., Overfelt, D., & Picou, J. S. (Eds.). (2007). *The Sociology of Katrina: Perspectives on a Modern Catastrophe.* Rowman & Littlefield.

Campanella, R. (2007). An Ethnic Geography of New Orleans. *Journal of American History, 94*, 704–715.

Chakraborty, J., Tobin, G. A., & Montz, B. E. (2005). Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review, 6*(1), 23–33.

Chang, L., & Lindell, M. K. (2005). *Hurricane Wind Risk Assessment for Miami-Dade County, Florida: A Consequence-Based Engineering (CBE) Methodology.*

CBE Institute Report. Mid-America Earthquake Center, University of Illinois at Urbana-Champaign.

Ciger, B. A. (2007). The "Big Questions" of Katrina and the 2005 Great Flood of New Orleans. *Public Administration Review, 67*, 64–76.

Collins, T. W., Grineski, S. E., & Aguilar, M. D. R. (2009). Vulnerability to Environmental Hazards in the Ciudad Juarez (Mexico)-El Paso (USA) Metropolis: A Model for Spatial Risk Assessment in Transnational Context. *Applied Geography, 29*(3), 448–461.

Colten, C. E. (2005). *An Unnatural Metropolis: Wrestling New Orleans from Nature*. Louisiana State University Press.

Costanza, R., Mitsch, W. J., & Day, J. W. (2006a). A New Vision for New Orleans and the Mississippi Delta: Applying Ecological Economics and Ecological Engineering. *Frontiers in Ecology and the Environment, 4*(9), 465–472.

Costanza, R., Mitsch, W. J., & Day, J. W. (2006b). Creating a Sustainable and Desirable New Orleans. *Ecological Engineering, 26*(4), 317–320.

Curtis, A., Mills, J. W., & Leitner, M. (2007). Katrina and Vulnerability: The Geography of Stress. *Journal of Health Care for the Poor and Underserved, 18*(2), 315–330.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards. *Social Science Quarterly, 84*(1), 242–261.

Cutter, S. L., & Emrich, C. T. (2006). Moral Hazard, Social Catastrophe: The Changing Face of Vulnerability Along the Hurricane Coasts. *Annals, American Academy of Political and Social Science, 604*, 102–112.

Cutter, S. L., Emrich, C. T., Mitchell, J. T., Boruff, B. J., Gall, M., Schmidtlein, M. C., et al. (2006). The Long Road Home: Race, Class, and Recovery from Hurricane Katrina. *Environment, 48*(2), 8–20.

Cutter, S. L., & Finch, C. (2008). Temporal and Spatial Changes in Social Vulnerability to Natural Hazards. *Proceedings National Academy of Sciences, 105*(7), 2301–2306.

Cutter, S. L., Mitchell, J. T., & Scott, M. S. (2000). Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographers, 90*(4), 713–737.

Dyson, M. E. (2006). *Come Hell or High Water: Hurricane Katrina and the Color of Disaster*. Basic Civitas Books.

Elder, K., Xirasagar, S., Miller, N., Bowen, S. A., Glober, S., & Piper, C. (2007). African Americans' Decision Not to Evacuate New Orleans Before Hurricane Katrina: A Qualitative Study. *American Journal of Public Health, 97*(12), 2112–2222.

Elliott, J. R., & Pais, J. (2006). Race, Class, and Hurricane Katrina: Social Differences in Human Responses to Disaster. *Social Science Research, 35*(2), 295–321.

Elliott, J. R., Sams-Abioudun, P., & Haney, T. (2007, April 12–14). *A City of Two Tales: Natural Disaster Neighborhood Differences in Response to Hurricane Katrina*. Paper presented at the Disaster and Migration: Hurricane Katrina's Effect on New Orleans.

Fekete, A. (2009). Validation of a Social Vulnerability Index in Context to River-Floods in Germany. *Natural Hazards and Earth System Sciences, 9*(2), 393–403.

FEMA (Federal Emergency Management Agency). (2005). *Hurricane Katrina Remote Sensing Data from NGA*. Retrieved March 20, 2009, from http://www.gismaps.fema.gov/2005pages/rsdkatrina.shtm

Fischetti, M. (2001). Drowning New Orleans. *Scientific American*, 76–85.

Frey, W. H., Singer, A., & Park, D. (2007). *Resettling New Orleans: The First Full Picture from the Census*. Brookings Institution Metropolitan Policy Program. Retrieved March 17, 2009, from http://www.frey-demographer.org/reports/Brook7-2ResettlingNewOrleans;theFirstFullPicturefromtheCensus.pdf

Fuchs, S. (2009). Susceptibility Versus Resilience to Mountain Hazards in Austria—Paradigms of Vulnerability Revisited. *Natural Hazards and Earth System Sciences, 9*(2), 337–352.

Fussell, E. (2007). Constructing New Orleans, Constructing Race: A Population History of New Orleans. *Journal of American History, 97*, 846–855.

Germany, K. B. (2007). The Politics of Poverty and History: Racial Inequality and the Long Prelude to Katrina. *Journal of American History, 97*, 743–751.

Greater New Orleans Community Data Center. (2008). *Census Population Estimates 2000–2007 for New Orleans MSA*. Retrieved March 17, 2009, from http://www.gnocdc.org/census_pop_estimates.html

Green, R., Bates, L. K., & Smyth, A. (2007). Impediments to Recovery in New Orleans' Upper and Lower Ninth Ward: One Year After Hurricane Katrina. *Disasters, 31*(4), 311–335.

Hartman, C., & Squires, G. D. (Eds.). (2006). *There Is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina*. Routledge.

Hewitt, K. (1997). *Regions of Risk: A Geographical Introduction to Disasters*. Longman.

Landry, C. E., Bin, O., Hindsley, P., Whitehead, J. C., & Wilson, K. (2007). Going Home: Evacuation-Migration Decisions of Hurricane Katrina Survivors. *Southern Economic Journal, 74*(2), 326–343.

Laska, S. (2008). What If Hurricane Ivan Had Not Missed New Orleans? *Sociological Inquiry, 78*(2), 174–178.

Laska, S., & Morrow, B. H. (2006). Social Vulnerabilities and Hurricane Katrina: An Unnatural Disaster in New Orleans. *Marine Technology Society Journal, 40*(4), 16–26.

Leong, K. J., Airriess, C., Chen, A. C.-C., Keith, V., Li, W., Wang, Y., et al. (2007a). From Invisibility to Hypervisibility: The Complexity of Race, Survival, and Resiliency for the Vietnamese-American Community in Eastern New

Orleans. In K. A. Bates & R. S. Swan (Eds.), *Through the Eye of Katrina: Social Justice in the United States* (pp. 169–185). Carolina Academic Press.

Leong, K. J., Airriess, C. A., Li, W., Chen, A. C.-C., & Keith, V. (2007b). Resilient History and the Rebuilding of a Community: The Vietnamese American Community in New Orleans East. *Journal of American History, 94*, 770–779.

Logan, J. R. (2006). *The Impact of Katrina: Race and Class in Storm-Damaged Neighborhoods*. Report by the American Communities Project Brown University. Retrieved March 17, 2009, from http://www.s4.brown.edu/katrina/report.pdf

Long, A. P. (2007). Poverty is the New Prostitution: Race, Poverty, and Public Housing in Post-Katrina New Orleans. *Journal of American History, 94*, 795–803.

Louisiana Public Health Institute. (2006). *2006 Louisiana Health and Population Survey*. Retrieved March 17, 2009, from http://popest.org/popestla2006/

Masozera, M., Bailey, M., & Kerchner, C. (2007). Distribution of Impacts of Natural Disasters Across Income Groups: A Case Study of New Orleans. *Ecological Economics, 63*(2–3), 299–306.

McCarthy, K., Peterson, D. J., Sastry, N., & Pollard, M. (2006). *The Repopulation of New Orleans After Hurricane Katrina*. Technical Report by the RAND Gulf States Policy Institute. Retrieved March 17, 2009, from http://www.rand.org/pubs/technical_reports/2006/RAND_TR369.pdf

Miller, D. S., & Rivera, J. D. (2007). Setting the Stage: Roots of Social Inequity and the Human Tragedy of Hurricane Katrina. In K. A. Bates & R.s. Swan (Eds.), Through the Eye of Katrina: Social Justice in the United States (pp. 15–35). : Carolina Academic Press.

Mock, N., Coates, J., Matthews-Sterling, M., & Papendieck, A. (2007, April 12–14). *Dynamics of Early Recovery in Two Historically Low Income Neighborhoods*. Paper Presented at the Disaster and Migration: Hurricane Katrina's effect on New Orleans.

Morrow, B. H. (1999). Identifying and Mapping Community Vulnerability. *Disasters, 23*(1), 1–18.

Myers, C. A., Slack, T., & Singlemann, J. (2008). Social Vulnerability and Migration in the Wake of Disaster: The Case of Hurricanes Katrina and Rita. *Population and Environment, 29*, 271–291.

Ngo, E. B. (2001). When Disasters and Age Collide: Reviewing Vulnerability of the Elderly. *Natural Hazards Review, 2*(2), 80–89.

O'Brien, P., & Mileti, D. (1992). Citizen Participation in Emergency Response Following the Loma Prieta Earthquake. *International Journal of Mass Emergencies and Disasters, 10*(1), 71–89.

Ortiz, E., & Plyer, A. (2008). *Valassis Lists Data as an Indicator of Population Recovery in New Orleans*. Retrieved March 17, 2009, from http://www.gnocdc.org/repopulation/ValassisTechDoc.pdf

Pastor, M., Bullard, R. D., Boyce, J. K., Fothergill, A., Morello-Frosch, R., & Wright, B. (2006). *In the Wake of the Storm: Environment, Disaster, and Race After Katrina*. Russell Sage Foundation.

Plyer, A. & Bonaguro, J. (2006). *Using U.S. Postal Service Delivery Statistics to Track the Repopulation of New Orleans and the Metropolitan Area*. Research Note by Greater New Orleans Community Data Center. Retrieved March 17, 2009, from http://www.gnocdc.org/USPS/research_note.pdf

Potter, H. (Ed.). (2007). *Racing the Storm: Racial Implications and Lessons Learned from Hurricane Katrina*. Lexington Books.

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008). A Sensitivity Analysis of the Social Vulnerability Index. *Risk Analysis, 28*(4), 1099–1114.

Tierney, K. (2006). Social Inequality, Hazards, and Disasters. In R. J. Daniels, D. F. Kettl, & H. Kunreuther (Eds.), *On Risk and Disaster* (pp. 109–127). University of Pennsylvania Press.

Tobin, G. A., & Ollenburger, J. C. (1993). *Natural Hazards and the Elderly*. University of Colorado, Natural Hazards Research and Applications Information Center.

Trujillo-Pagan, N. (2007). Katrina's Latinos: Vulnerability and Disasters in Relief and Recovery. In K. A. Bates & R. S. Swan (Eds.), *Through the Eye of Katrina: Social Justice in the United States* (pp. 147–168). Carolina Academic Press.

Wisner, B., Blaikie, P., Cannon, T., & Davis, I. (2004). *At Risk: Natural Hazards, People's Vulnerability and Disasters*. Routledge.

Correction to: Global Labour in Distress, Volume I

Pedro Goulart, Raul Ramos, and Gianluca Ferrittu

Correction to:

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6

The original version of the book was inadvertently published with incorrect affiliation. It has been corrected to read as below:

University of Bonn, Bonn, Germany
University of Warsaw, Warsaw, Poland
e-mail: ostark@uni-bonn.de

The updated version of these chapters can be found at
https://doi.org/10.1007/978-3-030-89258-6_4
https://doi.org/10.1007/978-3-030-89258-6_12

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6_29

Index[1]

[1] Note: Page numbers followed by 'n' refer to notes.

P. Goulart et al. (eds.), *Global Labour in Distress, Volume I*, Palgrave Readers in Economics,
https://doi.org/10.1007/978-3-030-89258-6

F

G

H