Giulia Mennillo
Thomas Schlenzig
Elmar Friedrich *Editors*

Balanced Growth

Finding Strategies for Sustainable Development

Springer

Management for Professionals

For further volumes:
http://www.springer.com/series/10101

Giulia Mennillo
Thomas Schlenzig
Elmar Friedrich
Editors

Balanced Growth

Finding Strategies for Sustainable Development

Springer

Editors
Giulia Mennillo
Institute of Economics
University of St. Gallen
St. Gallen, Switzerland

Thomas Schlenzig
Institute of Management
University of St. Gallen
St. Gallen, Switzerland

Elmar Friedrich
Institute for Economy
and the Environment
University of St. Gallen
St. Gallen, Switzerland

ISSN 2192-8096 e-ISSN 2192-810X
ISBN 978-3-642-24652-4 e-ISBN 978-3-642-24653-1
DOI 10.1007/978-3-642-24653-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011941774

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface by the Editors

Ever since the financial and economic crisis of 2008/09 and its accompanying spill-over effects, our globalizing world has uncompromisingly shown the flip side of its coin. Its crisis-prone character has intensified the discussion about our economic system's sustainability. Questions related to acceptable sovereign debt levels, suitable trade deficits and surpluses, and firms' growth targets have aroused high interest. In our opinion, this discussion must involve rethinking the qualitative and quantitative dimension of our present understanding of the nature of economic growth. The concept of balanced growth as a *way* of the global economy to grow might be one solution for the systemic risks and imbalances of our times.

The *kinds* of economic growth most countries aim for has created externalities with self-destructive traits concerning the system's ability to sustain itself. Is this inability based on the system requiring the exclusion of certain groups of countries, regions, and human beings? In other words, is inequality an embedded core element in our way of growing and achieving wealth? Exports and domestic consumption – the two prominent drivers of economic growth – are often also the origin of imbalance and instability. Excessive current account surpluses imply excessive current account deficits. The increase in global demand leads to shortages of supply and price increases, which may lead to social tensions. In an increasingly global and unevenly growing world, scientists and practitioners are deeply affected by these ambiguous developments. If the fault line lies within our society's *way* of growing, we encourage our reader to reflect on the unquestioned dogma of unlimited growth by taking the different elements of the debate into consideration.

We hope the reader of this book – accompanying this year's DocNet Management Symposium at the University of St.Gallen – obtains valuable insights from researchers and practitioners' theoretically and empirically grounded contributions. They are experts in the fields of macroeconomics, monetary economics, housing finance, strategic management, organizational control and governance, entrepreneurship, strategy consulting, human resources, and energy management.

We structured the book into three sections, according to different understandings of the concept of "balanced growth." Each section contributes to a formulation of perspective-related strategies. Section 1 highlights aspects related to the economics perspective. Section 2 deals with the business perspective on firms' growth strategies, while Section 3 focuses on questions related to the sustainable management of natural and human resources. Section 1 introduces the "growth imperative" – a concept based on the relationship between money creation and GDP growth. Furthermore, the role of interest rates and the direction of capital flows are highlighted concerning global imbalances, for example, in terms of per capita income differences. The related topic of housing bubbles with its interlinkages to the banking sector is examined across countries with different income levels. An analysis of the

positive impact of human capital on growth shows the differing effects between countries. Section 2 investigates strategies for firms' balanced growth. These chapters provide insights into the selection of the right growth mechanisms, as well as into the challenges of and success factors for sustainable firm growth, the growth strategies of sustainable entrepreneurs, and how firm board effectiveness can contribute to a firm's sustainable growth. Section 3 addresses the sustainable management of natural and human resources. Dwindling resources, rising global consumption, and a rapidly increasing demand for energy are among the main challenges for our generation. Not only natural resources are consumed at an unprecedented speed; we also stress the challenges faced by human resources, leading to record-breaking burnout statistics. In response to these problems, this section analyzes the opportunities of energy storage systems, offers insights into balanced HR management practices, and describes alternative solutions for the exploration and exploitation of resources in a case study on coffee production in sub-Saharan Africa.

This book is the product of many people's cooperation. We thank DocNet, the doctoral network at the University of St.Gallen, for providing us with the opportunity to organize the 9th DocNet Management Symposium, entitled *Balanced Growth: Finding Strategies for Sustainable Development* and to serve as the editors of this book. We also thank the authors and symposium speakers, some of whom authored chapters. All of them have time-consuming positions, but managed to provide contents that build and enrich this year's symposium book. Since the realization of such a project always requires much organizational and operational work, we thank Ilse Evertse and Johan Emerson Grobler for their commitment, which went far beyond the usual proofreading. Our thanks to Christian Rauscher and his Springer-Verlag team for guiding us through this work. We owe Philip Schuster and Larissa Zierow a debt of gratitude for their technical support.

Finally, we thank the various financial sponsors, including this year's lead sponsor, KPMG. In particular, we want to thank the Center for Organizational Competitiveness (RoCC) of the University of St.Gallen for supporting this book financially. RoCC is an interdisciplinary research center that hosts approximately 30 researchers and lecturers from different faculties. The Center analyzes the individual and collective skills of multidivisional company executives. Executives must satisfy the demands of different stakeholders such as investors, customers, competitors, and society at large in order to provide added value for their firms that exceeds single business units' abilities. The goal of the excellence initiative is to generate new knowledge to improve responsible strategic leadership.

In many contexts, terms like *sustainability* or *balanced growth* have unfortunately become buzzwords. In practice, they are often a desirable vision rather than an achievable objective. Why? Doubts may arise about the extent to which such concepts are compatible with a modern market economy. If this book can shed some light on how to realize a balanced way to grow and illustrates why a balanced growth path is a desirable strategy for policymakers, managers, employees, and other societal stakeholders, we would be more than thankful. We hope the reader discovers new insights and derives possible synergies from this interdisciplinary experiment.

Foreword by Prof. Dr. Müller-Stewens

University of St.Gallen

The major strategic leadership challenges may be addressing the core managerial dilemmas all companies face: Tension fields such as global standards vs. local needs, serving multiple stakeholder expectations, short-term vs. long-term interests, and so on. What is the right ratio to deal with each dilemma?

The answers are not as simple as they were a few years or decades ago. Companies and their contexts have changed dramatically and continue to do so. The global corporations of our time are increasingly complex. They face many more – often contradictory – constraints and conditions. Many companies now act in dozens of different political and cultural market settings, must manage a huge diversity of people and nationalities, and interact with a growing number of increasingly influential stakeholders (e.g., activists groups). These groups are able to pressurize companies, for example, by organizing their interests via social networks.

Leaders should be able to deal with and balance these increasing and diverse interests and expectations. They must understand the different underlying perspectives. They have to find ways to make sense of an often confusing complexity, in which decisions must often be taken without knowledge of all or even most of the possible – and often unintended – consequences and side effects.

Leaders often face situations where various options are available for decisions and, no matter which option they choose, they will do something wrong. Every option has its "price tag." Many options are inherently contradictory. For instance, in Saudi Arabia, seeking to uphold human rights imperatives may bring one into conflict with the cultures, customs, and practices of the country. Would you accept the wish of a female Arabian member of staff to take lunch with her male colleagues from Europe - even if it is against the law in some Arabian countries? There are no right or simple answers to such leadership dilemmas.

Leaders current face two primary dilemmas daily. First, how does one balance a short-term focus and a long-term focus on sustainable development? Some stakeholders demand short-term results, because they follow different strategic rationales. Companies must show progress in their quarterly reports. However, such efforts are not always sustainable. Second, how does one balance efficiency and growth? Both are major drivers of shareholder value. While companies must balance the logics of both efficiency and growth, there is not always a positive correlation between these two drivers. There is no easy way out of this dilemma; for many companies, balancing the two is often a challenging and confusing struggle.

In *The Age of Paradox*, Charles Handy writes: "*Part of the confusion stems from our pursuit of efficiency and economic growth, in the conviction that these are the necessary ingredients of progress. In the pursuit of these goals we can be tempted to*

forget that it is we, individual men and women, who should be the measure of all things, not made to measure for something else. It is easy to lose ourselves in efficiency, to treat efficiency as an end in itself and not a means to other ends."

The book you are now holding seeks to provide valuable insights into the above-mentioned challenges. The authors reflect on the different facets of *balanced growth*. The book also accompanies the interdisciplinary discussions at the DocNet Management Symposium 2011 to be held at the University of St.Gallen. The publication is kindly sponsored by the HSG research project 'Responsible Corporate Competitiveness' (RoCC).

Introductory remarks to the economics perspective on balanced growth

by Prof. Dr. Carl Christian von Weizsäcker

Max Planck Institute for Research on Collective Goods

Is balanced growth possible? And: If the answer is yes, is it desirable? Before we can answer these two questions, what is balanced growth and how does it relate to the concept of sustainability?

Binswanger is the only one of the five authors in the economics section to address this question: Why should there be growth at all? Summarizing his earlier work on this topic, he explains economic growth as a result of the institutional set-up of the capitalist financial system, which induces banks to create money. Contrary to neoclassical doctrine, he does not explain growth as a result of the interplay between individual utility maximizing behavior and profit maximizing behavior. The latter is derived from individuals' desire to improve their lot, from their "pursuit of happiness" – to cite the US Constitution. Binswanger strongly questions the desirability of boundless economic growth, specifically due to its effect on natural resources. Owing to my neoclassical background, I am not convinced by Binswanger's explanation that growth is due to money creation. Nevertheless, I share his concern about the depletion of natural resources. Is economic growth compatible with sustainability? And is the concept of balanced growth perhaps what people have in mind when they contemplate the potential compatibility of growth with the goal of sustainability?

In the four other papers, written by a much younger generation of researchers, the meaning of balanced growth is disconnected from the sustainability concept in the ecological sense of this word. Here, balanced growth basically means a movement in tandem of macroeconomic variables like GDP, real investment, human capital, and the like – perhaps also income distribution. All four papers are empirical and are concerned with phenomena thought to contradict "balanced growth." Thus, these phenomena are a sign that the actual worldwide economic growth is not balanced at all. The authors ask: What is the cause of the observed imbalances? The implicit norm is therefore that balanced growth is good. I return to this point when summarizing the results of the four papers.

Friedemann Roy is concerned with the recurrence of housing bubbles, which trigger financial crises. He explains this pattern as due to the ability of modern financial markets to provide large loans to families wanting to buy an own home. The housing sector is therefore highly leveraged. Depending on the macroeconomic conditions, valuation procedures may lead to an overvaluation of real estate property and, thus, to a surplus of new houses. This is a bubble, which then bursts and leads to a banking crisis. This crisis leaves remnants in the real economy for a

period of about six years. Clearly, we do not expect any of this to happen under conditions of balanced growth at all. Roy offers proposals that amount to a deleveraging of the housing sector.

Sunde and Vischer undertake an international cross-section analysis to answer the question: What influence does human capital have on economic growth? Their result is an interesting dichotomy. By means of a sophisticated econometric procedure, they endogenously form two groups of countries. One group shows that human capital makes an important contribution to economic growth, while the other shows no such influence. By and large, human capital has an influence on rich countries' economic growth, but not on poor countries' economic growth. What are the appropriate policy conclusions? I return to this question below.

Thomas Davoine is concerned with the net upward flow of capital from poor countries to rich countries. He discusses the hypothesis as investigated in the literature that this "mystery" is due to a large gap in capital market imperfections between rich countries (where the gap is small) and poor countries (where the gap is large). He is skeptical of economics' ability to contribute and improve the world in this respect. We do not yet know enough to give good advice. The upward flow of capital is implicitly considered to be contrary to the norm we might attach to the concept of balanced growth.

Daniel Kienzler is also concerned with international capital flows. His econometric analysis seeks to show that the large inflows of capital into the US had a significant impact on the country's long-term interest rates. This, he argues, also caused, or at least aggravated, the housing bubble, which itself re-enforced those capital imports. He also supports the hypothesis that the Fed's easy money policy in the early part of the last decade contributed to the housing bubble and, thus, to the ensuing financial crisis. If he is correct, we have another deviation from balanced growth.

Is it reasonable to accept "balanced growth" as a norm? This, I believe, should be one of the main topics of the DocNet Management Symposium 2011 in St.Gallen. The idea of economic growth originated from the observation of physical growth in nature, including human nature. Children grow up. Animals are born small and grow. Trees grow. But all individual growth comes to an end; furthermore, it is succeeded by decline and death. However, as a concept and an ideal, sustainability is the opposite of decline and death. Sustainability is connected with immortality. But mortals are not immortal. At best, societies of mortals may strive for immortality. Thus, the individual longing for immortality is transformed, i.e. socialized into a collective longing for collective immortality – a longing for "sustainability."

But then, can endless economic growth be balanced? Can it be compatible with sustainability? Economic growth, despite the connotation conjured by the term, cannot obtain its benchmark from individual growth of humans, of children, because it would then not only end, but would also gradually decline and die. Thus, this benchmark would contradict sustainability. However, it would also be problematic to use the form of individual growth, say that of children, as a benchmark of economic growth. Healthy children's growth is actually balanced. The way the

different limbs grow is balanced, although limbs do not grow proportionally. But, if we abandon children's growth as a benchmark, it is no longer clear why balanced growth should be the norm.

Economic growth as we know it is strongly driven by what we know as "innovation." Schumpeter's concept is strongly linked with his idea of "creative destruction." Neither destruction nor creativity is associated with the idea of balance. Nobody would speak of "balanced destruction" or "balanced creativity." Thus, although innovation is the main source of growth, it is – almost by definition – something that destroys "balance." Can the aggregate picture of the economy be one of "balanced growth"? I have my doubts.

In terms of economic growth, two of the most pressing concerns are the question of distributive justice across different nations and the problem of climate change due to the emission of greenhouse gases as a byproduct of human production processes. We obviously need to address these concerns.

Table of Contents

Part A

Balanced Growth – the economics perspective

Growth imperative and money creation – a new outlook on growth dynamics

Hans Christoph Binswanger

Modern economic development is characterized by a permanent tendency towards growth. I will argue that growth depends substantially on the continuing and boundless creation of money and the resulting dynamics.

With regard to money, one should first know what money is today. Money is anything with which you can pay. Today, one can pay with paper money from the Central Bank, which is the legal tender, but also with bank deposits, which are the result of loans by the banks. These loans are credited on the liability side of the banks' balance sheets. Payment with bank deposits is either made by transfer order or by credit card. Currently, approximately 95% of all payments are made with bank deposits and only 5% with paper money (and coins). Deposits can be exchanged for paper money at banks, but these can no longer be converted into gold at the Central Bank as in the past. The obligation to convert was definitely abolished in the 1970s. Since then, the Central Bank can issue any amount of paper money, regardless of the gold reserves. Consequently, the amount of money – bank deposits and paper money – can be increased from year to year. This is correctly described as money creation. Without the golden fetter of the past, this can carry on indefinitely.

Banks produce money. They produce it by granting loans to private firms, the government, and households. It is 100% *new* money, because no equivalent amount is debited from another account. A small part of this – about 5% – will actually be redeemed in paper money. Banks must therefore have sufficient reserves to meet this obligation. They can, however, always refinance themselves through the Central Bank, i.e. obtain new paper money by selling financial assets, mainly government debts. The Central Bank can, I would like to repeat, continuously issue new paper money, because there is no obligation to exchange this for gold.

A decisive factor is the increase in the money supply which occurs through the increase in loans (credits) to private firms and leads to economic growth, because firms normally use the money for investments, i.e. to increase production capacities, which leads to an increase in goods. The newly created money is therefore converted – not into gold anymore, but into produced goods. In this way, the real gross domestic product (GDP) increases on the basis of money creation.

*

Growth has developed into a perpetual dynamic process, which sustains itself in that it creates the precondition that makes continuous growth possible. How does this happen?

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_1,

Firms must pay for the labor supply provided by households and for other inputs *when* they are used, i.e. when production starts. Firms can, however, sell products only after they have been produced. Production takes time. The firms therefore require an advance – money not yet earned – to pay for labor and other inputs. This advance is the firm's capital. The capital includes borrowed capital (external funds), mainly bank loans, and equity capital (internal funds), which households make available to the firms – currently mostly by buying shares –, or which is retained profits. Without preliminary financing, i.e. without the conversion of money into capital and without money advances, there is no market, no investments, and no economic growth.

These money advances require profit. Why? Because investing capital is risky. The firms selling products in the future do not know what portion of the money spent on investment will flow back when they will have sold their products in the future. The future is always uncertain. Therefore firms expect a profit that will not only cover the interest rate to be paid on the borrowed capital, but will also leave sufficient net profit so that the equity capital can compensate for the risk.

For the functioning of the economy as a whole, the prospect of a profit should therefore always be higher than the prospect of a loss. The expected profit should hence be positive. This is only the case if the probability of profits was and remains generally higher than the probability of losses, i.e. the sum of all firms' profits should exceed the losses.

How can such a positive profit balance in the aggregate arise? Firms' profits are defined by their earnings less their expenses, or, more precisely, by the difference between their earnings and their expenses for the inputs necessary for the production of the goods from which they obtain their earnings – thus, the difference between earnings and costs. In order to make profits in the aggregate, aggregate earnings must always be larger than aggregate expenditure. How can this happen? How can all market participants earn more than they spend? It is obviously not possible if the same amount of money just goes around in circles, i.e. if the money that firms pay households for their labor supply is simply used by the households to buy products that the firms have produced with their help. Under these circumstances, firms' earnings would always be equal to their expenditure. The sum of profits and losses would then not have a positive profit balance. Therefore, a positive profit balance can, on the whole, only be achieved if there is a constant inflow of money.

But, how does money flow into the modern economy? We already know this: By firms obtaining loans from banks, which the banks can always provide through the process of money creation, i.e. by increasing the money supply through the credit channel. Firms use the borrowed money, together with their reinvested net profit, to purchase additional labor and other inputs. The incomes of households as providers of the labor supply therefore grow with the increase in GDP, while firms' profits increase with the households' growing incomes, since the households spend their income on the products that firms have produced.

The point is that households immediately spend the money which they receive for providing labor to firms. Thus, household expenses immediately become the earnings of the firms, which sell them their products. However, firms can only sell

products that have already been produced previously, i.e. before they received new money created by loans from the banks and which they paid to the households in exchange for providing labor. Therefore firms' earnings increase before the production costs of the goods they sell increase. The result is positive aggregate profits, provided the economy continues to grow. Hence, the circular flow of money becomes a growth spiral perpetuating itself. Owing to money creation, output growth is the necessary condition for profits. In turn, profits create the precondition for money to be used as capital, and these profits thus enable further growth (see Figure 1).

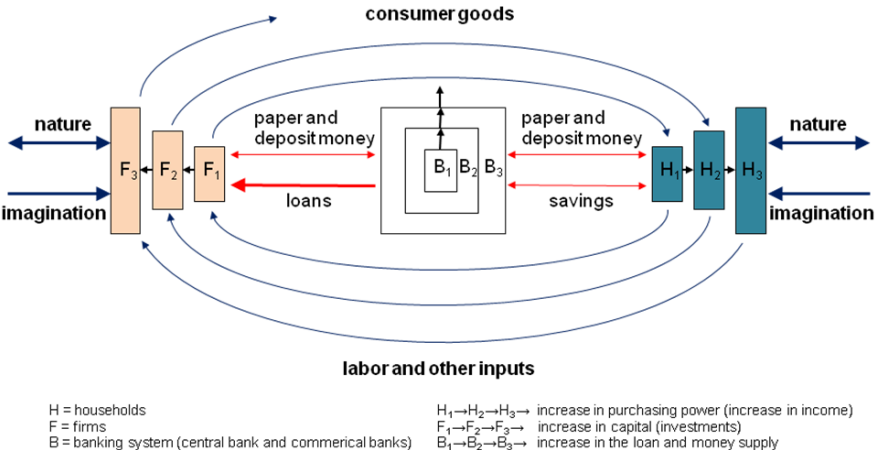


Figure 1: The growth spiral

The circular flow occurs between households and firms – firms buy households' labor as an input factor for production, and households buy goods produced by the firms. The flow expands with each "rotation," due to a) the firms' investments and b) the increase in the households' income. This creates a growth spiral. The banking sector effects the payment transactions between the households and the firms and expands due to the growth of the banks' balance sheets. The expansion of the growth spiral's scale is due to a) money creation through the credit channel, b) the extraction of natural resources (with a simultaneous release of waste and emissions into nature), and c) human imagination inventing new products and methods of production.

*

The drawback, however, is that economic growth leads to a growth imperative. The growth process *must* always continue; without a continuous expansion of the money supply due to the need to finance new investments, which triggers additional demand, the subsequent increase in supply due to the preceding investment activity will fall into a void. The increase in demand is necessary to compensate for the already occurred increase in supply. If this did not happen, the effect of the production increase of the previous period's investments would occur without being absorbed by the income effect of a new investment. The profit margin would decrease correspondingly. If the profit margin falls below the minimum level that firms or investors expect in return for their investment risk, firms will no longer

provide replacement investments and will gradually stop production. Eventually, interests can then no longer be paid. An increasing number of firms will make losses and, by going bankrupt, will cease production activity. Instead of growing, the economy begins to shrink. The growth spiral inverts and becomes a shrinking spiral. This results in a growth *imperative* in the sense that under a certain benchmark of a minimum growth rate, the alternative to growing is shrinking. In other words, stability and zero growth are not possible in the modern economy. Given the conditions of our monetary system, growth is essential for the system (see Figure 2).

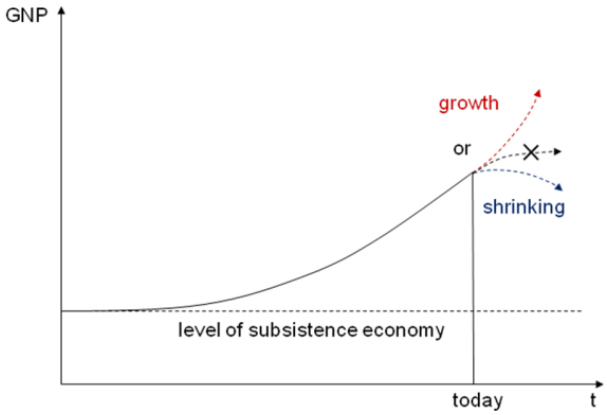


Figure 2: The growth imperative

However, this does not necessarily mean that the growth rate has to be arbitrarily high. A minimal growth rate is sufficient. Basically, this rate is determined by the credit interest rate that banks charge firms, banks' capital ratios, and the net return investors expect in order to take on the risk of investment. Based on plausible assumptions, I have calculated the minimum level of the global (!) growth rate as 1.8% in my book *Die Wachstumsspirale* [The growth spiral]. This figure could even be lower, if systemic risk is reduced.

*

The prerequisite for maintaining growth is, however, that there are no stumbling blocks to hamper it. This, on the other hand, cannot be guaranteed. Growth is jeopardized by crises that become increasingly acute the more advanced the growth dynamic becomes.

The danger of financial and economic crises like the ones in 2007/08 is of prime importance. Primarily, these crises are the result of an excessive money creation which does not finance real production, but leads to speculative purchases of financial assets with the expectation of ever-increasing asset prices due to a constant money supply increase. When the expected future price increase is higher than the interest rate, loans are made to purchase assets in order to make quick profits. Such speculation is, however, not entirely risk-free, since the interest rates could at a cer-

tain point increase. This occurs if the Central Banks fear inflationary pressures due to the increasing issuing of speculative loans, which causes an inflated money supply. In order to prevent this, Central Banks increase the interest rates. Since too high interest rates reverse the leverage effect of speculative loans, a financial crisis results. This happened in 2008.

But what would happen if there were no financial crises? Would everything then be satisfactory? No, because the tendency towards economic growth can only be fed if there are sufficient available natural resources to obtain raw materials and energy, which provide the basis for increased output. Economic growth, however, is increasingly confronted with long-term environmental depletion. This does not only refer to natural resources, but also to limits to waste and emissions, because the world, i.e. nature, is finite. Its exploitation cannot be arbitrarily prolonged and extended. The foreseeable shortage, especially of energy, certain raw materials, and food, had already led to price increases before the last crisis. These rising prices are again in the spotlight, as they increase the inflation risk besides exacerbating the environmental crisis.

*

The analysis of the prerequisites for economic growth indicates where reforms to control growth must primarily occur – with the *reform of the monetary system*. The idea is to generate money in a different way than through escalating and unchecked money creation by commercial banks. The starting point can be the 100% money concept of the American economist Irving Fisher, the most important American economist of the 20th century. He developed a concept after the 1929 crisis, which is now more topical than ever. According to this proposal, the Central Bank has the exclusive right to money creation; banks are obliged to cover the deposits with 100% reserves. This happens through the banks' deposits at the Central Bank, which can be always converted into bank notes. The Central Bank – and only the Central Bank – is thus autonomously and solely in charge of money creation. The Central Bank is given the possibility to proactively and not just, as at present, reactively determine the extent of money creation so that possible crises and collateral damage due to growth are prevented as far as possible. The aim is to avoid:

- inflation and deflation of commodity prices
- speculative inflation of financial assets, and
- unlimited economic growth beyond the current capacity of a finite world.

The legitimization of strengthening the role of the Central Bank and the duty to do so results from the following argument: Without state legislation, which ensures that the paper money issued by the Central Bank is the legal and definitive tender – that is, it can no longer be converted into gold –, there would only be a small fraction of the currently circulating and constantly increasing money supply.

The reform of the monetary system must be complemented by a *reform of corporate law*. Specifically, revision is required of legislation regarding joint stock firms. Owing to their legal structure, these firms are aimed at unlimited growth and

cannot pursue a strategy of moderation. This is due to the share value at the stock exchange being based on the sum of an infinite series of discounted expected profits. Moreover, since expected profits are higher the higher the current investments are, the faster a firm grows. For this reason, shareholders are willing to accept smaller dividends in exchange for larger profits in the future. The smaller dividends are overcompensated by the share price increase! Since this applies to all corporations, which are clearly the economy's major players, the tendency towards unlimited growth is preprogrammed for the entire economy. What can be done?

Public firms are legal entities and, thus, a societal invention they would not exist without legislation. Consequently, shareholders' rights are dependent on legislation. It is therefore possible to change legislative content with reform to avoid crises caused by excessive growth. This could include the creation of two categories of shares as was partially done in the past: registered shares [in German: *Namensaktien*] and bearer shares [in German: *Inhaberaktien*]. Registered shares would not be traded on the stock exchange, and over-the-counter (OTC) selling would only be allowed after three years. They should keep an indefinite duration. Bearer shares might, however, be traded on stock exchanges, but would only have a duration of 20 or 30 years, with a redemption of the initial investment at the end of their validity term. Such a limitation automatically reduces the unlimited increase in share values and, at the same time, limits the scale of possible declines. This would substantially reduce the risk of repetitive financial and economic bubbles.

For long-term investments, other forms of organizations should be further promoted: e.g., foundations and cooperatives.

A foundation is oriented towards a particular goal and not only towards the realization of profits. The founders can receive a pension funded by the profits, but most of the profits will be reinvested in the foundation in line with its original goal. The foundation's assets are not traded on the stock exchange. This means that the foundation is not subject to speculation. The foundation may acquire additional autonomy through the incorporation of a limited liability firm, which performs the operational part of the business.

Another possibility is to enhance the cooperative concept. The cooperative is designed as a mutually supportive group and is thus more member-related. In principle, the profits are retained in the business. Each cooperative member has an equal share in the cooperative's registered capital and an equal voting right. A member can usually only recoup his money if he finds a new cooperative member to take his place. As in the case of the foundation, the shares of the cooperative are not traded on the stock exchange. This means that the cooperative, too, is not subject to speculation.

These reforms in the area of monetary policy and corporate law would not only significantly avert crises in the economy, but would also lessen the demand for natural resources and dampen the economic system's environmental impact. The demand and impact would not become zero, because growth – albeit on a lower level – will continue. Further measures, such as new regulations and the tightening of existing ones, are therefore required for more energy and resource efficiency. All these measures will, however, be insufficient if the global GNP growth remains on

the recent high level – with the exception of the 2007/08 crisis – of 4-5%. Deceleration of growth is necessary to ensure that the increase in efficiency per product *unit* is not always offset by an increase in the *number* of products, with the ultimate aim to realize sustainable economic development.

References

Binswanger, H. C. (2009). *Vorwärts zur Mässigung*. Hamburg: Murmann Verlag.
Binswanger, H. C. (2006). *Die Wachstumsspirale*. Marburg: Metropolis-Verlag.

Beyond balanced growth: The effect of human capital on economic growth reconsidered

Uwe Sunde and Thomas Vischer

Abstract:

Human capital plays a central role in theoretical models of endogenous growth. Nevertheless, the existing empirical evidence on human capital's effect on economic growth is mixed. Part of this weak evidence might be the result of overly restrictive specifications of the estimated empirical growth models. A more flexible estimation framework's results reveal that human capital does indeed have a strong positive effect on economic growth, but this effect differs substantially across countries. In particular, the evidence suggests that human capital has a stronger effect in countries in which the population faces favorable living conditions in terms of life expectancy and geographic characteristics, while the effects do not differ substantially between countries that differ with regard to their political institutions' quality.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_2,

Table of Contents

1 Introduction

Knowledge and the ability to invent and innovate are considered the key determinants of economic development. This productive resource, which is embodied in people, is often referred to as *human capital.* This terminology reflects that human capital must be acquired through a process of formal or informal education. Since human capital is particularly important for economies with few other resources, education policy receives particular attention. In virtually all countries of the world, education-related matters are currently high on policy makers' agenda. Simultaneously, an increasing body of academic literature examines the determinants of education outcomes and the most effective ways to use the resources invested in the education sector.

Nevertheless, empirical estimates of human capital's growth effect have delivered mixed results. This lack of consensus among economists about human capital's empirical effect on growth appears surprising in the light of human capital's central role not only in the public debate, but also in many models of endogenous growth.[1]

A closer look reveals that most estimates have been obtained with empirical models that make restrictive assumptions about the specification of the empirical growth framework that they use for estimation. On the one hand, existing empirical studies are either based on estimation frameworks that restrict attention to the dynamics along the balanced growth path (such as studies in the tradition of Mankiw, Romer, and Weil, 1992) [2], or are based on those that adopt a growth accounting framework (for instance, studies along the lines of Benhabib and Spiegel, 1994), which assumes identical technology across countries. On the other hand, existing empirical studies have concentrated exclusively on one of various channels through which human capital is thought to affect growth. Human capital can either be seen as a factor of production, similar to physical capital or raw labor, or as the prerequisite for adopting new technologies or for innovation. In the former, human capital growth affects the output growth potential, while in the latter, the size of the existing human capital stock determines the output growth potential.

The ongoing research on human capital's empirical growth effects has recently addressed these shortcomings. Owen et al. (2009) investigated whether the assumption of a balanced growth path for all countries is appropriate in specifications of empirical growth models. Using a more flexible estimation framework that allows for different growth processes for distinct groups of countries, these authors' results suggest different growth regimes across the world. However, their results also suggest that the level of human capital might have a negative growth effect in some countries. Sunde and Vischer (2011) investigated a different source of misspecifi-

[1] In endogenous growth models, long-term economic growth is explained by the variables within the model. This is in contrast to exogenous growth models, where long-term growth is explained by factors, such as technological progress, that are determined outside the models.

[2] In macroeconomics, a balanced growth path refers to a situation in which major aggregates have the same growth rate over time. In the neoclassical Solow model, for example, the considered aggregates are output and the capital stock (Temple, 2008).

cation, due to the restriction to only one channel through which human capital might affect growth. They show that if both channels are empirically relevant, then estimates that restrict attention to only one channel lead to serious misspecification and thus deliver biased estimates of human capital's growth effect. Sunde and Vischer's empirical results from replications of existing studies with an appropriately extended specification reveal human capital's quantitatively and statistically significant positive growth effect through both channels.

However, it is still unclear whether human capital matters for growth in the same way across countries when both channels – the production factor and the prerequisite to innovate – are taken into account. An investigation of this question requires the combination of a flexible estimation framework, such as the one used by Owen et al. (2009), with an extended growth specification that accounts for both human capital channels, as suggested by Sunde and Vischer (2011).

The remainder of this chapter presents the results from such an analysis. Section 2 contains a description of the data and the empirical approach. Section 3 describes the results. The findings suggest that human capital is particularly important in a group of countries where human capital has a positive effect on growth through both of the channels discussed in the literature. This group of countries appears to be more developed, with better geographic features and a higher life expectancy. Human capital's growth effect is substantially smaller and statistically insignificant in the other group of countries. This finding has far-reaching implications, which are discussed in the concluding section (Section 4).

2 Empirical model and data

2.1 Data

The data for the analysis span the period 1965 to 2005 and are taken from different sources. The data on the GDP per capita and investment are from the Penn World Tables (version 6.3). The population data and the data on life expectancy at birth originate from the UN Population Division. The data on human capital correspond to the working age population's (aged 25-65) average years of schooling, as gathered by Barro and Lee (2010). The quality of political institutions is measured as an index taken from Freedom House and has been used in recent empirical studies (see, e.g., Acemoglu et al., 2008).

The estimation is conducted in long differences from 1965 to 2005, based on a sample of 94 countries for which all variables are available in the data sources. Table 1 contains the data's summary statistics.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Average annual growth in GDP per capita 1965-2005	0.0179	0.0159	-0.0148	0.0627	94
Change in log mean school years (25+) 1965-2005	0.0256	0.0147	0.0045	0.0718	94
Log mean school years (25+) in 1965	0.8176	0.9875	-1.9348	2.3245	94
Change in log capital per worker 1965-2005	0.0423	0.029	-0.0336	0.1533	94
Log GDP per capita 1965	8.1322	0.9871	6.2286	9.945	94
Log life expectancy 1965	4.0038	0.2107	3.537	4.3022	94
Absolute latitude	24.0278	16.6057	1	64	94
Freedom House Index 1965	0.6507	0.2877	0.1	1	82

2.2 Empirical model

In order to account for the possibility that countries might differ in their growth processes and, hence, in the extent to which human capital affects economic growth, the analysis follows Owen et al. (2009) and adopts a data-driven strategy. Rather than grouping countries ad hoc, for example, according to their income level at some pre-specified date, the estimation allows for different latent growth processes whose prevalence and structure are estimated from the data.

This is done by estimating the growth process by maximum likelihood, using finite mixture models. These models pre-specify the growth process as a function of the main determinants of growth in light of the theoretical predictions. In addition to human capital, these determinants include the initial income level in order to account for convergence dynamics, as well as the investment rate, which reflects the change in physical capital. But rather than assuming that the respective coefficients are the same for all countries in the sample, the finite mixture approach allows for classifying countries into a pre-specified number of groups by means of the similarity in their growth rates' conditional distributions. A distinct set of coefficients is estimated for the growth determinants in each of these groups. Furthermore, the estimation provides an estimate of a particular data point's (country observation) propensity to belong to one of these groups. The propensity of belonging to one group can be estimated as a latent variable. Alternatively, this propensity can be specified as depending on particular characteristics, for example, the life expectancy, geographic features, or institutions. The reader is referred to Owen et al. (2009) for a detailed technical description of the finite mixture approach in the context of empirical growth models.

3 Empirical results

3.1 Baseline results

The empirical results are presented in Table 2. As a benchmark, Column 1 contains the results of the empirical growth model's estimation, which constrains the coefficients to be the same across all countries in the pooled sample. Human capital has a significant positive effect on growth through the change in human capital over the observation period, as well as through the level of human capital at the beginning of the observation period in 1965. An increase of 1% in the working age population's average schooling years increases growth by 0.56%. The accumulation of physical capital also has a positive effect on growth, with an elasticity of 0.22% increase in the growth rate for a 1% increase in the capital stock. Finally, the results suggest conditional convergence, as indicated by the negative coefficient for the log GDP per capita in 1965. These results essentially replicate results presented by Sunde and Vischer (2011) and point towards human capital's positive effects through both channels, which are individually and jointly significant. Nevertheless, these results still maintain the assumption of a balanced growth path for all countries in the sample, with human and physical capital having identical effects on growth across all countries.

Columns 2 and 3 relax this assumption and present the results of a more flexible growth process estimation by applying a finite mixture model with two latent growth regimes. This specification treats the propensity to be in one of two different growth regimes as a latent variable, which is estimated jointly with the coefficients for the explanatory variables (human capital changes and levels, physical capital investments, and convergence) by maximum likelihood.

The estimates deliver an endogenous split of the sample countries into two distinct regimes: Regime 1 with 61 countries, and Regime 2 with 33 countries. A comparison of the estimation results across the two regimes reveals a striking disparity in the coefficient estimates. In Regime 1, investment in physical capital is the only explanatory variable with a (marginally) statistically significant positive effect on growth. Neither human capital, nor the convergence term, displays any effect. In contrast, in Regime 2, all the explanatory variables have a strong effect. Human capital affects growth through both channels by means of positive and significant coefficients that are even slightly larger than those in the pooled specification in Column 1. The accumulation of physical capital also positively affects growth with a slightly larger coefficient than in the pooled sample. Finally, the convergence term is about 1.5 times larger.

These results suggest that the growth process might follow a different path in different countries where the role of growth determinants varies across different growth regimes. However, the unconditional endogenous split into two regimes raises the question as to which factors are actually behind these different growth regimes.

Table 2: Baseline results: Standard regression models and finite mixture models

	Dependent variable: Average annual growth rate of GDP per capita 1965-2005		
	All	**Regime 1**	**Regime 2**
Change in log mean school years (25+) 1965-2005	0.5587** [0.260]	0.3277 [0.9132]	0.5846** [0.2757]
Log mean school years (25+) in 1965	0.0181*** [0.004]	0.0126 [0.0114]	0.0185** [0.0085]
Change in log capital per worker 1965-2005	0.2222*** [0.053]	0.1911* [0.0989]	0.2473*** [0.0439]
Log GDP per capita 1965	-0.007*** [0.002]	-0.001 [0.0026]	-0.010*** [0.0024]
Constant	0.0324* [0.017]	-0.0065 [0.0409]	0.0658*** [0.0166]
Regime membership:			
Constant		0.3141 [1.1917]	
Observations	94		
Countries in class		61	33
Bayesian information criterion		-521.58	

Notes: All regressions in long-difference specifications with one observation per country over the period 1965-2005. Column 1: All displays results from *ordinary least squares*. Columns 2 and 3 (*Regime* 1, *Regime* 2) display the results from the estimation of a finite mixture model. Robust standard errors are provided in brackets. ***, **, * indicate significance at the 1-, 5-, and 10-percent level, respectively.

More precise knowledge of the characteristics that lead countries to be characterized by a growth process along the lines of Regime 1 or Regime 2 would be particularly valuable for the question at hand. On the one hand, if it were known that a particular country is more likely to be characterized by Regime 2, the case for policy interventions targeting the economy's education structure would be much stronger than if the country were better characterized by Regime 1. On the other hand, if it were known that the regime propensity could be affected by policy, for example, because the propensity to exhibit a growth process as in Regime 2 would be increased if institutional quality were improved, this would indicate potentially important complementarities between policy interventions in the education sector and institutional quality.

3.2 Taking a closer look

The following analysis attempts to shed light on the country characteristics that potentially explain the classification of countries into different growth regimes. The analysis considers three potential determinants of a country's growth regime: absolute latitude, life expectancy, and the quality of the political institutions. The choice of these variables is related to an ongoing debate in the growth literature on the fundamental causes of long-term development. This debate has specifically focused on the question as to whether geographic features, which are largely exogenous, or institutions – i.e., human-made rules for the conduct of social, political, and economic interactions – are the main fundamental cause of long-run growth. Latitude is a proxy for geographic features, such as climate and agricultural conditions, the extrinsic disease environment, and related features that might affect a country's growth and development potential. Latitude has the advantage of being a readily available measure of relevant geographic features, and has therefore been frequently used in empirical work. Life expectancy is somewhere intermediate in the debate on geography vs. institutions, since life expectancy is strongly affected by the extrinsic climate and disease conditions, which relate to the geographic location. Simultaneously, however, institutional quality also strongly influences life expectancy, for example, in terms of the health system's efficiency and coverage, as well as its financing, which usually involves substantial degrees of redistribution. Finally, the quality of political institutions, as reflected by an index of political liberties, represents humanly devised institutions that are clearly endogenous and malleable in the long-run.

Table 3 presents the results of two specifications, each of which allows for two distinct growth regimes, similar to Columns 2 and 3 in Table 2. Each of these specifications includes a particular classification variable.

Columns 1 and 2 present the results of the estimation of a finite mixture model, which stipulates two distinct growth regimes, with a country's absolute latitude as the main determinant of its propensity to be characterized by a particular regime. This estimation reveals an estimated regime split with 69 countries in Regime 1 and 25 countries in Regime 2. Countries with a lower latitude (located closer to the equator) are more likely to be characterized by the Regime 1 growth process and countries with a larger latitude by the Regime 2 growth process. Given that a larger latitude is typically associated with a higher level of economic development, this implies that Regime 1 essentially represents less developed countries, whereas Regime 2 represents more developed countries. This specification's fit is better than that of the specification with a constant as the latent regime membership, as estimated in Table 2 and as reflected by the Bayesian information criterion's lower value of -542.27 compared to -521.58.

Table 3: Finite mixture models with regime determinants: Latitude and life expectancy

	Dependent variable: Average annual growth rate of GDP per capita 1965-2005			
	Regime 1	**Regime 2**	**Regime 1**	**Regime 2**
Change in log mean school years (25+) 1965-2005	0.3482 [0.3312]	0.7734*** [0.1830]	0.1674 [0.403]	1.0131*** [0.372]
Log mean school years (25+) in 1965	0.0148** [0.0059]	0.0225*** [0.0019]	0.0047 [0.008]	0.0240*** [0.006]
Change in log capital per worker 1965-2005	0.1896*** [0.0553]	0.0663 [0.0935]	0.1181** [0.053]	0.2166*** [0.060]
Log GDP per capita 1965	-0.0051** [0.0026]	-0.016*** [0.0031]	-0.0067 [0.006]	-0.009*** [0.002]
Constant	0.0287 [0.0194]	0.1209*** [0.0335]	0.0444 [0.036]	0.0419** [0.019]
Regime membership:				
Absolute latitude	-0.1478** [0.0620]			
Log life expectancy 1965			-16.9*** [6.233]	
Observations				
Countries in class	69	25	34	60
Bayesian information criterion	-542.27		-538.52	

Notes: All regressions in long-difference specifications with one observation per country over the period 1965-2005. Columns 1 and 2 (*Regime* 1, *Regime* 2) display the results from a finite mixture model with absolute latitude as the regime classification variable. Columns 3 and 4 (*Regime* 1, *Regime* 2) display the results from a finite mixture model with log life expectancy in 1965 as the regime classification variable. Robust standard errors are provided in brackets. ***, **, * indicate significance at the 1-, 5-, and 10-percent level, respectively.

In terms of growth determinants, the results show that changes in human capital have no growth effect in Regime 1, whereas the initial stock of human capital has a positive but comparably small growth effect. In contrast, in Regime 2, the change in human capital and its initial level have large and statistically positive growth effects. In contrast, physical capital accumulation exhibits a significant effect in Regime 1 countries, whereas it exhibits no growth effect in Regime 2. The conditional convergence process appears to be stronger in Regime 2 countries. Finally, the larger constant indicates that baseline growth is more rapid in Regime 2 countries.

In total, these results suggest that human capital is important for growth, but that the degree to which human capital matters for growth differs across countries. Furthermore, the potential to adopt and operate technologies – as reflected by the initial stock of human capital's effect – as well as the increase in the available pool of human capital, affects growth much stronger in Regime 2. Most importantly, however, the growth processes seem to differ strongly across countries.

Columns 3 and 4 present results of the estimation of a model with life expectancy as a determinant of regime classification. This estimation delivers an estimated regime classification of 34 countries into Regime 1 and 60 countries into Regime 2, with countries with a high life expectancy more likely to be classified as Regime 2. Since countries with higher life expectancy are considered more developed, this again implies that the classification reflects a split into less developed countries (Regime 1) and more developed countries (Regime 2). Compared to the results obtained with latitude as the regime classification variable, life expectancy leads to a substantially different classification of countries. Nevertheless, the goodness of fit is comparable, or slightly worse, with a Bayesian information criterion value of -538.52.

In terms of the determinants of growth, the estimates again display substantial differences in the coefficients between the two regimes. Human capital appears not to have any effect on growth in Regime 1, neither in terms of the change in human capital, nor in terms of the initial stock. In contrast, human capital has a strong positive effect on growth in the countries classified as Regime 2 countries. Both effects are substantially larger in size than in the benchmark estimates obtained with the pooled sample reported in Column 1 of Table 2. Physical capital accumulation has a comparable positive growth effect in both regimes. The conditional convergence process and the constant growth rate are only significant in Regime 2. In short, these results corroborate the previous findings of substantial heterogeneity in the growth process across countries. More importantly, the results suggest that human capital is an important determinant of growth through both channels, but only for countries that appear to be more developed in the sense of exhibiting a higher life expectancy at the beginning of the observation period in 1965.

Table 4 presents results of institutions' quality as the regime predictor variable. Owing to missing values, the sample size is reduced to 82 countries in this specification. The estimation is conducted by using a binary regime quality variable as the regime predictor, which takes the value 1 for all countries with a quality of democracy index above the median in the sample, and zero otherwise.[3] The results deliver a classification of 29 countries into Regime 1 and 53 countries into Regime 2, where countries with a better institutional quality are more likely to be classified as Regime 1. In contrast to the previous results, Regime 1 represents countries with better political institutions, which usually refers to more developed countries, whereas Regime 2 represents countries with political institutions of a lower quality.

[3] Alternative specifications using a binary variable that codes countries above and below the mean index, or directly using the respective index values for the regime classification, deliver qualitatively identical results.

However, institutional quality does not appear to be a strong predictor of regime membership, as is indicated by the Freedom House indicator's insignificant effect. With a Bayesian information criterion value of -465.81, this model's fit is even worse than that of a model with an unspecified latent regime membership propensity, as reflected in Columns 2 and 3 in Table 2.

4 Conclusion

The estimation results of finite mixture models with different regime classification determinants all deliver a broadly coherent picture. Human capital has a strong effect on growth through both channels, as a production factor and as a prerequisite for the adoption of new and more efficient technologies. The results thereby corroborate Sunde and Vischer's (2011) recent findings. These findings suggest that specifications accounting for only one of these channels might deliver biased estimates of human capital's growth effect. The results also show that these effects are not homogeneous across countries. Rather, it appears that human capital is particularly important for growth in countries with a comparably high level of development, which is proxied by better geographic conditions in terms of absolute latitude or higher life expectancy. According to these findings, and in line with the conclusions by Owen et al. (2009), a balanced growth path does not appear to be an appropriate characterization of growth from a comparative development perspective. The empirical results do not, however, provide a definite answer with regard to the question as to whether geography or institutions is the fundamental determinant of growth. Rather, the results are consistent with both geography and institutions playing some role in shaping the growth process. Nevertheless, the results also indicate that the heterogeneity in the growth process, particularly in terms of human capital's growth effect, is more pronounced in subsamples that differ in terms of geographic latitude or life expectancy than in terms of institutional quality.

The findings also provide a new aspect to the debate on poverty traps and conditional convergence. From a more applied, policy-oriented perspective, the results suggest that, concerning education policies, developed countries might have a higher return in terms of growth than underdeveloped countries. This implies that improvements in institutional quality might produce a double dividend: they might have additional indirect effects by making other production factors, such as human capital, more effective with regard to fostering economic growth.

Table 4: Finite mixture models with regime determinants: Political institutions

	Dependent variable: Average annual growth rate of GDP per capita 1965-2005	
	Regime 1	**Regime 2**
Change in log mean school years (25+) 1965-2005	-0.2082 [0.2172]	0.5211 [0.4379]
Log mean school years (25+) in 1965	0.0133*** [0.0047]	0.0165** [0.0069]
Change in log capital per worker 1965-2005	0.1437*** [0.0366]	0.3002*** [0.0575]
Log GDP per capita 1965	-0.008*** [0.0014]	-0.007*** [0.0025]
Constant	0.0648*** [0.0111]	0.0393* [0.0221]
Regime membership:		
Freedom House 1965	2.1679 [2.1508]	
Observations		
Countries in class	29	53
Bayesian information criterion	-465.81	

Notes: All regressions in long-difference specifications with one observation per country over the period 1965-2005. Columns 1 and 2 (*Regime* 1, *Regime* 2) display the results from a finite mixture model with a binary indicator taking the value 0/1 for countries with a Freedom House index 1965 below/above the median as the regime classification variable. Robust standard errors are provided in brackets. ***, **, * indicate significance at the 1-, 5-, and 10-percent level, respectively.

References

Acemoglu, D., Johnson, S., J. Robinson, & Yared P. (2008). Income and Democracy. *American Economic Review*, 98(3), 808-842.

Barro, R.J., & Lee, J.W. (2010). A New Data Set of Education Attainment in the World, 1950-2010. *NBER Working Paper*, 15902.

Benhabib, J., & Spiegel, M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics,* 34(1), 143-173.

Mankiw, G.N., Romer, P., & Weil, D. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107, 407-437.

Owen, A.L., Videras, J., & Davis, L. (2009). Do all countries follow the same growth process?. *Journal of Economic Growth*, 14(4), 265-286.

Sunde, U., & Vischer, T. (2011). Human Capital and Growth: Specification Matters. *University of St.Gallen, Dept. of Economics Discussion Paper No. 2011-31.*

Temple, J. (2008). "Balanced Growth". *The New Palgrave Dictionary of Economics.* Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008.

Global imbalances and long-term interest rates

Daniel Kienzler

Abstract:

The bond yield conundrum describes the unusual behavior of US long-term interest rates between 2004 and 2007. These rates' constancy or decline was considered puzzling, as US monetary policy was restrictive during this period. Theory and historical evidence suggest that such restrictiveness would cause an increase in long-term interest rates. Global imbalances in the form of sustained and increasing international capital flows to the US are believed to be a promising explanatory factor for the interest rate development during this period. A cointegration approach is applied to assess international capital movements' influence on US long-term interest rates. The inclusion of international capital flows in the analysis can indeed explain a considerable part of the low long-term interest rate environment, which played a decisive role in the recent financial and economic crisis.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*,
Management for Professionals, DOI 10.1007/978-3-642-24653-1_3,

Table of Contents

1 Introduction

In the United States (US), long-term nominal interest rates behaved unusually between 2004 and 2007. When the US central bank – the Federal Reserve – raised short-term interest rates, long-term interest rates remained constant or even decreased. In contrast, historical evidence shows that longer-term interest rates typically move in the same direction as short-term interest rates. From a theoretical perspective, the expectations hypothesis of the term structure of interest rates also suggests, other things being equal, increasing long-term rates when short-term rates increase. The then Federal Reserve Chairman, Alan Greenspan, called the behavior of long-term interest rates during this period "a conundrum."

This phenomenon attracted much attention because long-term interest rates are the main link between monetary policy and net investment, which, in turn, is a main component for economic growth on the macro level (Case et al., 1999). The low long-term interest rate environment is considered a causing factor of the asset market bubble that led to the financial and economic crisis that started in 2007, which caused severe output losses in developed countries and critical growth decelerations in developing countries (Obstfeld and Rogoff, 2009).

Owing to continuing and increasing global current and capital account imbalances at the beginning of the 2000s, researchers considered international capital flows into US financial markets a possible explanatory factor for the so-called bond yield conundrum. Long-term interest rates could have been depressed by a long-lasting relative increase in demand for long-term bonds over supply, induced by foreign purchases of these securities (Frey and Moëc, 2006).

This chapter examines the link between global imbalances and long-term interest rates by focusing on the bond-yield conundrum period. To better understand why long rates behaved puzzlingly during this period, the following section provides a theoretical background to interest rate determinants. Section 3 analyzes the actual development of US long-term interest rates in recent years. Subsequently, a notion is developed of global imbalances and international capital flows, and international capital flows' possible influence on US long rates is described. Based on the previous analysis, Section 5 presents an econometric investigation of the relationship between capital flows and long-term interest rates. Section 6 presents the argument that international capital flows' impact on long-term interest rates contributed considerably to the recent financial and economic crisis. Section 7 summarizes the findings.

2 Long-term interest rates – theory

According to Goodfriend (1998), there are two basic theories for understanding long-term interest rates: The Fisher relation and the expectations hypothesis. The former states that real interest rates (i.e. the rate of return on capital) and expected inflation determine nominal interest rates. The latter establishes a relationship

between long-term interest rates and the current, as well as expected, short-term interest rates.

2.1 The Fisher relation

The Fisher (1896) relation represents an equilibrium condition established by arbitrage: If inflation expectations increase, the expected real return on nominal bonds will decrease relative to the real return on capital. As a result, investors sell bonds, thereby depressing the bond price and increasing the bond yield until the expected real return on financial capital and real capital are equal. With R as the gross nominal return (i.e. one plus the nominal interest rate), r as the gross real return (i.e. one plus the real interest rate), and π^e as the expected inflation rate, this relation can be formally expressed as $R = r\,\pi^e$ (Ireland, 1996).

The Fisher relation is mainly criticized for disregarding economic agents' risk aversion (Wolman, 2006). This issue can be addressed in the following consumption-based asset pricing framework that derives from Lucas (1978). Suppose there is a representative consumer with an infinite planning horizon and that a single consumption good exists in the economy. The consumer's utility function reads

$$v = E_t \sum_{j=0}^{\infty} \beta^j \ln(c_{t+j}) \tag{1}$$

where c_{t+j} is consumption in period $t+j$ and β is a discount factor taking on values between zero and one. E_t is the expectation operator, given information known at time t. The economy is further described by the following characteristics (Ireland 1996): There are two types of assets, real bonds and nominal bonds. A real bond costs the investor one unit of the consumption good in period t and has a gross real return of r_t units of the consumption good in period $t+1$. The number of real bonds the investor buys in t is b_t. A nominal bond costs the investor one US dollar (USD) in t and gives her a gross nominal return of R_t USD in $t+1$. The amount of nominal bonds the investor purchases in t is B_t. The price of the consumption good in t is P_t USD and the representative agent receives y_t units of the consumption good as income in t (that is, y is real income). With these specifications, the following budget constraint can be established for period t in real terms:

$$c_t + b_t + \frac{B_t}{P_t} \leq y_t + r_{t-1} b_{t-1} + \frac{R_{t-1} B_{t-1}}{P_t}. \tag{2}$$

The investor's sources of funds in t – denoted on the right-hand side of (2) – consist of her income in t, the payoff from maturing real bonds she purchased in t-1, and the real payoff from maturing nominal bonds she bought in t-1. The investor's expenditure on consumption, real bonds, and nominal bonds (in real terms), which are denoted on the left-hand side of equation (2), may not exceed her income. In this setting, the representative agent wants to maximize (1) subject to (2) by choosing c_t, b_t and B_t in each period t.

Merging this intertemporal maximization problem's first-order conditions and applying the covariance formula $E(XY) = Cov(X,Y)+E(X)E(Y)$ for general X and Y yields the equilibrium condition

$$R_t^{-1} = r_t^{-1} E_t(\pi_{t+1}^{-1}) + \beta Cov_t(g_{t+1}^{-1}, \pi_{t+1}^{-1}). \quad (3)$$

$g_{t+1}=(c_{t+1}/c_t)$ is the gross rate of consumption growth and also the ratio of the marginal utilities in t and $t+1$, respectively. $\pi_{t+1}= (P_{t+1}/P_t)$ denotes the gross rate of inflation between t and $t+1$. Furthermore, the real interest rate in this model is equal to the marginal rate of intertemporal substitution, expressed as $E_t(g_{t+1}/\beta)$.[1] Put differently, the real interest rate compensates the investor for postponing consumption to the future and depends on economic agents' time preference (Allsopp and Glyn, 1999).

Equation (3) is a generalization of the Fisher relation: The nominal interest rate depends on the ex ante real interest rate and expected inflation. Additionally, however, the conditional covariance between expected consumption growth and the expected inflation rate enters as a determinant of the nominal interest rate. This component can be interpreted as a risk premium (Crowder and Hoffman, 1996): If the covariance term is positive, this implies that a high (low) expected real return on nominal bonds coincides with a high (low) expected marginal utility of consumption. In other words, the investor expects to receive a high (low) return – measured in consumption good units – on her nominal bond exactly when consumption is expected to be of high (low) value to her. This would increase the demand for nominal bonds and, hence, lower the nominal interest rate. Vice versa, if the covariance term is negative, the demand for nominal bonds would decrease and nominal interest rates would increase (Evans and Wachtel, 1990).

2.2 The expectations hypothesis

Similar to the Fisher relation, the expectations hypothesis is established by arbitrage considerations. Suppose economic agents face a situation allowing them to invest a certain amount of money in a long-term coupon bond, or in a sequence of short-term coupon bonds, whose aggregated time to maturity equals the long-term bond's time to maturity. If investors compare the return on both investment opportunities and perceive, for example, that the return on the long-term bond is higher than that on the sequence of short-term bonds, they will rebalance their portfolios in favor of long-term bonds. The long-term bond price will thus increase and the long-term yield will simultaneously decrease. Conversely, the short-term bond price will decrease and the short-term yield will rise. These capital movements continue until the returns on both investment strategies are equal because investors will then have no incentives to shift between different maturities. In other words, the long-term interest rate is the geometric mean of the current and expected future short-term interest rates (Bofinger, 2001).

[1] The marginal rate of intertemporal substitution expresses the rate at which an economic agent is willing to give up one unit of consumption today in order to obtain a unit of consumption tomorrow.

The expectations hypothesis implicitly posits that short-term and long-term securities are perfect substitutes and that short-term forward rates[2] are unbiased estimates of expected short-term interest rates. That is, if the forward rate for a short-term investment that becomes effective in a specific future period were higher than the expected short-term interest rate for this period, economic agents would augment their forward contract purchases. This would increase these contracts' price and depress their yield until the forward rate equaled the expected short-term interest rate (Spahn, 2006). Following this reasoning, one could easily obtain long-term interest rates by examining forward rates.

However, Hicks (1946) points out that risk-averse creditors are assumed to require compensation for the liquidity loss they incur by lending money. As a bond's time to maturity increases, the investment's liquidity decreases, increasing the risk that lenders will have to sell the investment at a capital loss. Hence, creditors are interested in shorter-term lending, whereas debtors prefer long-term loans to ensure higher planning reliability. In such a situation, borrowers are willing to pay a higher interest rate to convince lenders to grant them longer-term loans and enter the forward market. In contrast, lenders require a higher interest rate to enter the forward market to compensate them for facing the risk of capital losses. Accordingly, the forward rate consists of two components: The expected short-term interest rate and a risk or term premium. It is this risk premium that renders forward rates biased estimates of expected future interest rates. This modified expectations hypothesis therefore posits that long-term interest rates not only reflect movements in future short-term interest rates, but also vary with risk premia. Only if the latter are constant can movements in long rates be attributed to changes in short rates.

3 The bond yield conundrum

The actual development of long-term interest rates, represented as the interest rate on 10 year US treasury bonds, is shown in Figure 1 (solid line). A remarkable downward trend can be observed since the 1980s. Long rates dropped from more than 14% in 1981 to almost 3.5% in 2003 and have subsequently remained at low levels. The data suggest that long-term interest rates' defining factors, as identified in the previous section – namely inflation expectations, real interest rates, and risk premia – have contributed to this downward trend. For example, as reported by Bandholz et al. (2009), expected inflation – measured by core inflation – has trended significantly downward from 14% to around 2% since 1980, the largest decrease occurring during the first half of the 1980s.[3] Hauner and Kumar (2006) and Alexopoulou

[2] Forward rates are interest rates agreed upon today for transactions in some future period.

[3] Core inflation (i.e. the annual percentage change in the consumer price index excluding energy and food prices) can be regarded as the persistent part of inflation, since the excluded food and energy prices are deemed very volatile. Core inflation thus reflects the underlying inflationary pressure in an economy and is used to measure longer-term inflation expectations (Landau, 2000).

et al. (2006) regard declining inflation expectations as the main reason for declining long yields. Gruber and Kamin (2009) report that, after surging at the beginning of the 1980s, US real interest rates have also declined since the mid-1980s, although less than inflation expectations. Lower risk premia are often associated with lower macroeconomic variability, which implies less risk of incurring capital losses, resulting in investors demanding a lower compensation for this risk. Rudebusch et al. (2006) report that US GDP growth's volatility fell from 5 percentage points in the mid-1980s to approximately 1 percentage point in 2006, and that inflation volatility also decreased over this period, but to a lesser extent. Lower short-term interest rates, represented by the central bank policy rate in Figure 1 (dashed line), reflect the aforementioned macroeconomic variables' downward trend.

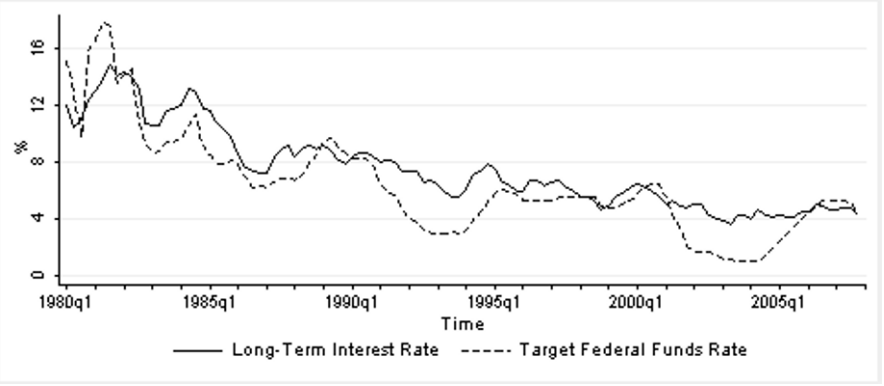


Figure 1: Long-term interest rate in the US and target federal funds rate since the 1980s (source: Federal Reserve Bank of St. Louis)

Although only sketchy and descriptive, these examinations indicate that long rates' overall development can be understood fairly well by looking at the theoretically derived determinants of nominal long-term interest rates. However, this does not apply to 2004 to 2007. In this period, the development of long-term interest rates in the US became the focus of many researchers and practitioners because long-term interest rates behaved puzzlingly regarding both theoretical and empirical considerations. In 2004, the Federal Open Market Committee began to tighten monetary policy. It raised the target federal funds rate from 1.25% in the second quarter of 2004 to 3.25% in the second quarter of 2005; nevertheless, long-term interest rates fell by approximately 0.4 percentage points. This development was confusing because theory – particularly the expectations hypothesis outlined in Section 2.1 – and past empirical evidence suggest that long rates should actually increase when short rates increase. Alan Greenspan, who was then chairman of the Federal Reserve, expressed his surprise in a speech in early 2005:

> "[...] long-term interest rates have trended lower in recent months even as the Federal Reserve has raised the level of the target federal funds rate [...]. This development contrasts with most experience, which suggests that, other things being equal, increasing short-term interest rates are normally accompanied by a rise in longer-term yields. The simple mathematics of the yield curve governs the relationship between short- and long-term interest rates. [...] For the moment, the broadly unanticipated behavior of [...] bond markets remains a conundrum."

After Greenspan's statement, long-term interest rates continued to act unusually. From the third quarter of 2005 to the fourth quarter of 2006, the target federal funds rate was again increased by 2 percentage points, but long-term interest rates stayed roughly the same. Brière et al. (2006) note that this long yield behavior was even more puzzling as US growth continued and oil prices were rising constantly during the conundrum period. The US budget deficit was also increasing, which meant that the worsening debt position should have boosted long-term interest rates (Hauner and Kumar, 2006).

The conundrum period is evident in Figure 1. When the target federal funds rate rose quite sharply from mid-2004 onwards, the long-term interest rate first decreased somewhat and was more or less at the same level at the end of the tightening period as it had been at the beginning. It is also apparent from Figure 1 that, in the conundrum period, long rates' development differed strikingly from that of long rates in all previous monetary policy tightening periods. In 1980-1981, 1983-1984, 1988-1989, 1993-1995 and 1999-2000, long-term interest rates increased without exception when the target federal funds rate was raised.

4 Global imbalances and US asset markets

Several researchers – such as Bandholz et al. (2009), Brière et al. (2006), and Alexopoulou et al. (2006) – suggest that international capital flows into treasury bonds are an explanatory factor for the bond yield conundrum. If economic agents believe that international capital flows will last for a longer period, they will perceive altered demand and supply conditions for bonds and demand a lower interest rate because securities can then be sold more easily to international investors (Frey and Moëc, 2005).

Bernanke (2005) introduced the notion that a global savings glut could cause international capital flows into US securities, which became a widely debated issue. This is the view that if particular world regions experience an interconnected change in their savings-investment relationship, this has lasting influences on US financial markets.

4.1 Savings, investments, and capital flows

The interrelationship between savings, investments, and international capital flows is best illustrated by simple macroeconomic accounting (e.g., Frenkel and John, 2007). National income Y is created by production for consumption C and invest-

ment I, and by net exports of goods and services to foreign countries CA; therefore, $Y = C + I + CA$. For simplicity, it is assumed that the balance of trade equals the current account, so that net exports (imports) comply with a current account surplus (deficit) of the same magnitude. Equally, national income can be expressed as $Y = C + S$, where S represents total savings in an economy. Equating the two expressions for Y, one obtains $S = I + CA$. This equation can be further broken down into $(S_{pr} + S_{go}) = (I_{pr} + I_{go}) + CA$ by dividing savings into private (pr) and government (go) savings as well as investment into private and government investment. This expression can be rearranged to $S_{pr} = I_{pr} + (I_{go} - S_{go}) + CA$. The term $(I_{go} - S_{go})$ can be viewed as the government budget deficit BD. The I/S identity for an open economy then reads $S_{pr} = I_{pr} + BD + CA$.

Owing to double-entry bookkeeping in the balance of payments, a current account surplus (deficit) matches a capital account deficit (surplus) and, thus, net capital exports (imports). Hence, private savings are used to finance private investment, the government budget deficit, and net capital exports (loans to foreign countries). The last equation clarifies that when a country's national private savings exceed its national private investments and the government budget deficit, the remainder of the savings flows abroad to finance investment or government deficits there; the country would therefore experience net capital outflows and a current account surplus. Conversely, if national private savings fall short of national private investment and the government budget deficit, a country will experience a current account deficit and, consequently, net capital inflows.

4.2 A savings glut perspective of global imbalances

If various regions or countries exhibit excess savings relative to investment (and budget deficits), they will all exhibit current account surpluses, as well as capital account deficits, which is the notion behind the global savings glut.

Bernanke (2005) believes the global savings glut is primarily rooted in currency and financial crises in Asia during the 1990s, which resulted in severe economic distortions. Basically, many East Asian countries experienced large capital inflows until the mid-1990s and, hence, current account deficits. Implicit government guarantees of private investments attracted additional capital inflows, leading to unduly risky capital allocations. The bursting of the ensuing asset price bubbles led to huge losses for firms and financial intermediaries, who could no longer cover their (often short-term) liabilities. In addition, implicit government guarantees were no longer sustainable, having become too expensive (Krugman, 1998). These developments resulted in a financial market crisis that triggered large capital outflows from East Asian countries, a strong devaluation of regional currencies, and – in the end – a sharp decline in economic activity (Corsetti et al., 1999).

In crises like the East Asian one, Bernanke (2005) sees a reversal effect regarding international capital flows. Owing to the financial and currency crises, uncertainty prevailed and economic agents revised their consumption behavior, thereby increasing savings. In contrast, investment opportunities shrank. This scenario led to current account surpluses in not only immediately affected countries, but also in

other countries – most importantly China – that feared the impacts of similar future crises.

If exchange rates do not float freely, current account surpluses imply a build-up of foreign exchange reserves through central bank foreign currency purchases. In turn, the associated improvement in price competitiveness boosts current account surpluses. Such an official capital export was particularly attractive for crisis-ridden countries since a stimulation of the economy via export-led growth was a good opportunity to mitigate unemployment problems. Furthermore, current account surpluses and increasing foreign exchange reserves provided a buffer against potentially destabilizing capital outflows, which often occur during currency crises.

In essence, during the conundrum period, developing countries evolved from capital importers to large capital exporters. This development is against common economic reasoning, which would suggest capital flows from developed to developing countries, because the latter exhibit lower capital-to-labor ratios and, hence, higher marginal products of capital.[4] Figure 2 demonstrates the abovementioned development concerning developing Asian economies. The dashed line shows these countries' current account balance. It is evident that they exhibited current account deficits and capital imports until the mid-1990s. In 1997, the current account deficit turned into a surplus, which grew to nearly 400 billion USD (left graph) or approximately 7% of GDP (right graph) in 2008.

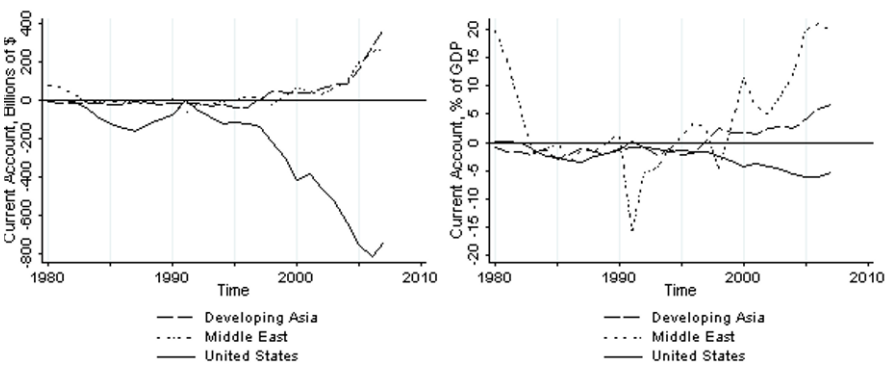


Figure 2: Current account balances of selected country groups (source: International Monetary Fund, World Economic Outlook Database)

Clarida (2007) has another explanation, especially since 2004, for the global savings glut: the rising oil price, which has led to oil exporting countries, many of which are located in the Middle East, earning huge profits. As can be seen in Figure 2's right-hand graph, the sharp increase in oil prices boosted Middle East countries' current account surpluses up to around 20% of GDP (dotted line). Figure 2 also clarifies the huge deterioration in the US current account since the mid-1990s, which worsened after 2001 and only picked up slightly in 2007 (solid line).

[4] For a more in-depth discussion of the direction of capital flows between developed and developing countries, see the chapter written by Thomas Davoine.

The left-hand graph in Figure 2 specifically suggests a connection between the steadily growing current account surpluses of Asia and the Middle East's developing countries and the increasing US current account deficit. However, the mere observation of excess savings relative to investments and, hence, that current account surpluses and capital outflows are prevalent in developing countries around the world does not justify direct implications for the US current account. Some observers even argue that the US current account is simply the result of too few US savings relative to investments due to declining private savings and, particularly, a high government budget deficit. If this hypothesis were true, one should have seen increasing real interest rates in the US since the mid-1990s due to a high demand for savings relative to supply (Dooley et al., 2005). However, as observed in Section 3, the real interest rate fell over this period. This militates against the idea that purely domestic factors were responsible for capital flows into the US and suggests a connection between the current account surpluses in developing countries and the current account deficits in the US.

4.3 The link between capital flows and US asset markets

According to Bernanke (2005), the US asset market is the channel through which this relationship between the two phenomena could work. Since developing countries' excess savings flowed to the US, the stock market experienced a huge upswing, leading to two effects: First, US consumers experienced a wealth effect induced by rising stock prices, which led them to consume more and increase their demand for imports. This effect was intensified by the higher purchasing power of the appreciating USD. Second, higher profit opportunities increased the demand for investments, also promoting the need for foreign financing when national savings declined. After the stock market downturn in 2000, investment needs declined, while the savings glut persisted, manifesting in declining real interest rates. Since 2004, the "recycling" of petrodollars – that is, the re-channeling of oil revenues into the US capital market – seems to have played a major role in declining real interest rates and risk premia (Clarida, 2007). The resulting increased demand for houses due to low mortgage rates brought about rising house prices, which again promoted more consumption, lower savings, and contributed to a further deterioration in the US current account.

The remaining question is why the savings glut "unloaded" specifically in the US and less so in other industrialized countries. A widely held view is that crisis-ridden developing countries are incapable of supplying sufficient financial assets to absorb domestic savings; they therefore demand saving instruments from industrialized countries (Caballero et al., 2008). Owing to US financial markets' depth and sophistication, they are particularly attractive (Legg et al., 2007). Furthermore, the USD's status as the major reserve currency might also have contributed to international capital flows affecting mainly US interest rates (Idier et al., 2007).

Since this work is concerned with treasury yields, Figure 3 depicts net foreign capital flows into US treasury securities over four quarters, scaled by GDP (left graph). There is a sharp increase in foreign net purchases of US treasuries in 2001,

reaching a peak in mid-2004 when foreign capital inflows into treasuries amount to 3.5% of GDP. Figure 3's right-hand graph shows that foreign holdings of US treasury securities increased sharply from 2001 to 2004, and have remained at high levels (more than 15% of GDP) since then.

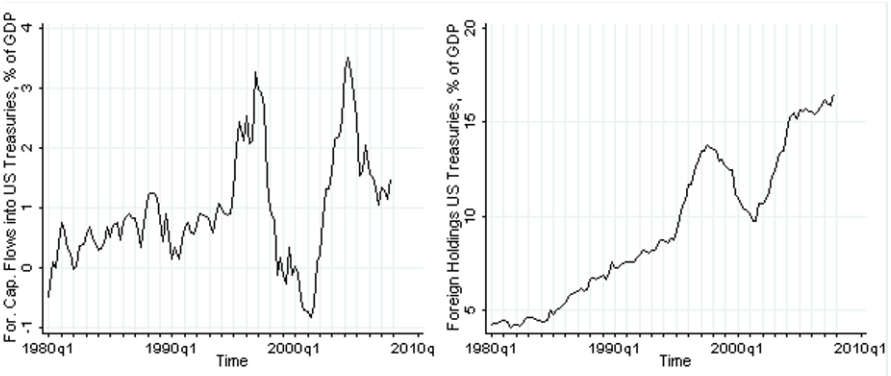


Figure 3: Foreign capital flows into US treasuries and foreign holdings of US treasuries (source: Federal Reserve Flow of Funds Account)

5 Econometric analysis

The above analysis suggests that international capital flows play an important role in explaining the bond yield conundrum, as they are very likely to have had a dampening impact on long-term interest rates. In the following, the role of international capital movements with regard to unusually low US long rates is therefore examined in an error-correction model (ECM). Only an intuitive description is provided of the econometric procedure to ensure the reader finds the analysis comprehensible. Readers interested in the technical details of ECMs are referred to Banerjee et al. (1993).

First, the relevant data are described, followed by a brief explanation of the econometric method and the estimation results.

5.1 Data

As outlined in Section 2.1, in a simple version of a consumption-based asset pricing model, the level of nominal long-term interest rates depends on three factors: the real interest rate, which, in turn, depends on expected consumption growth, inflation expectations, and a risk premium, which reflects the conditional covariance between expected consumption growth and the expected inflation rate. Owing to the variables' considerable persistence, expected consumption growth can be represented by the one-year percentage change in real per capita consumption, while inflation expectations can be measured as the one-year percentage change in the consumer price index (CPI) omitting food and energy (i.e. core inflation). As before, the interest rate on 10 year US treasury bonds (constant maturity) is used to

represent long-term interest rates. All these data are obtained from the Federal Reserve Bank of St. Louis. The risk premium is calculated as the rolling 8-quarter empirical covariance between expected consumption growth and expected inflation (i.e. the last 8 data points of the respective series – including the current ones – are used to calculate the covariance for each quarter). As in Rudebusch et al. (2006), foreign capital flows into US treasuries, measured as changes in foreign holdings over four quarters and scaled by GDP, are obtained from the Federal Reserve's Flow of Funds Account.

Quarterly data from the first quarter of 1983 to the fourth quarter of 2007 are used for the subsequent estimation procedures. The chosen period reflects three considerations: First, data from the 1980s onward ensure a reasonably large data set. Second, by choosing 1983 as the beginning of the data set, difficulties arising from monetary policy regime changes, which could have exerted uncontrolled influences on long-term interest rates, are circumvented. Third, the time period encompasses the bond yield conundrum.

5.2 Methodology

Applying a Dickey-Fuller test to the series of long-term interest rates, real interest rates, inflation expectations, and risk premia reveals that the null hypothesis of non-stationarity cannot be rejected on a 1% significance level. In addition, a Dickey-Fuller test of the first differences of the same variables rejects the null hypothesis, indicating that the variables are I(1), i.e. integrated of order one. A Johansen (1988) maximum eigenvalue test shows that, among the considered variables, there is one independent cointegration relationship.

If two or more nonstationary variables are cointegrated, a linear combination of these variables is stationary. The basic notion behind cointegration is that the relevant variables form an equilibrium relationship. Consequently, when the described system is in disequilibrium, the resulting disequilibrium error provides useful information about the variable of interest's future movement. It should increase if the disequilibrium error is negative and vice versa. If such an equilibrium relationship actually exists, the disequilibrium error is stationary (Banerjee et al., 1993). The data's described properties allow for applying the Engle and Granger (1987) two-stage procedure to estimate an ECM. In the first stage, a cointegration equation is estimated and the estimated disequilibrium errors are obtained. In the second stage, changes in the long-term interest rate are regressed on first differences of the first-stage explanatory variables, on lags of the latter, possibly on additional explanatory variables, and on the lagged estimated disequilibrium error (the so-called error-correction term); this is the ECM. Note that the second stage involves only stationary variables.

5.3 Results

As a first step, a model should be estimated of long-term interest rates, the specification of which is motivated by the theoretical considerations in Section 2.1. This model is used as a baseline approach representing the "traditional" influences on

bond yields. As a second step, the model is augmented by international capital flows into treasury bonds. An assessment can then be made regarding whether these capital movements have a statistically and economically significant influence on long-term interest rates.

Table 1 presents OLS estimation results for the purely domestic model. The first panel of Table 1 contains the estimated coefficients of the cointegration equation.[5] Examining the coefficients, it is clear that higher expected inflation (*exin*) and higher expected consumption growth (*excg*) translate, as expected, into a higher nominal long-term interest rate (*ltir*). With regard to the risk premium (*rp*) coefficient, the negative sign is also in line with the theoretical considerations outlined in Section 2.1.

Table 1: Estimation results – baseline equilibrium variables

Cointegration equation (dependent variable: $ltir_t$)							
$exin_t$	$excg_t$	rp_t	*constant*				
1.678	0.591	-0.329	0.001				
ECM (dependent variable: $\Delta ltir_t$): domestic variables							
ect_{t-1}	$\Delta ltir_{t-1}$	$\Delta exin_t$	$\Delta excg_t$	$\Delta excg_{t-1}$	$\Delta excg_{t-2}$	*constant*	
-0.241	0.396	0.381	0.151	0.106	-0.117	-0.041	
***	***	**	**	*	*		
(-4.41)	(4.08)	(2.08)	(2.29)	(1.73)	(-1.91)	(-1.03)	
ECM (dependent variable: $\Delta ltir_t$): international capital flows incorporated							
ect_{t-1}	$\Delta ltir_{t-1}$	$\Delta exin_t$	$\Delta excg_t$	$\Delta excg_{t-1}$	$\Delta excg_{t-2}$	$\Delta flow_t$	*constant*
-0.224	0.371	0.382	0.158	0.096	-0.131	-0.228	-0.040
***	***	**	**		**	**	
(-4.16)	(3.90)	(2.13)	(2.45)	(1.58)	(-2.17)	(-2.34)	(-1.04)

*,** and *** indicate significance on a 10%, 5%, and 1% level, respectively.

The associated ECM is specified according to a general-to-specific approach. The analysis started with a broad model comprising two lags of first differences of each variable and contemporaneous first differences of *exin*, *excg*, and *rp*. This general model was then gradually reduced on the basis of t-tests and F-tests. Table 1's second panel shows the estimated coefficients and the t-statistics (in parentheses) of the resulting model. Δ indicates a change. The adjusted coefficient of determination for this domestic model is 0.277.[6] The lagged change in the long-term interest rate is highly significant, with the expected positive sign, while changes in inflation expectations positively affect changes in long rates during the same period. Contemporaneous and first lag changes in consumption growth expectations positively affect

[5] Note that no standard errors or t-ratios are reported for the estimated coefficients, because although the OLS estimates are super-consistent in the case of cointegration, the distributions of the coefficient estimates are highly nonstandard (Stock, 1987).

[6] This is not an unusually low value for the adjusted coefficient of determination in an ECM.

changes in long rates. The second lag of the same variable affects the dependent variable negatively, but only marginally significantly. The error-correction term is highly significant. Its sign is negative, and therefore in accordance with the theoretical considerations outlined above. If the lagged disequilibrium error takes on a positive (negative) value, a downward (upward) movement is induced towards the equilibrium value.

Table 1's third panel shows the estimation results when first differences of international capital flows (*flow*) are incorporated into the otherwise unchanged ECM. The coefficients of the other variables hardly change at all, which is a desirable result from a robustness perspective. Compared to the previous model, the adjusted coefficient of determination increases to 0.311. International capital flows enter the model significantly and with the expected sign. According to the estimation, a 1 percentage point increase in the change of total foreign purchases results, other things being equal, in a 0.228 percentage point decrease in the change in long-term interest rates.

While the above results present evidence of international capital flows affecting long-term interest rates's dynamic behavior, it would be interesting if such a relationship could also be established as an equilibrium in a cointegration sense. Some researchers have therefore suggested that an international capital flow variable should be inserted into the long-run cointegration relationship between the levels of interest rates and the levels of explanatory variables (e.g. Frey and Moëc). This results in the estimates displayed in Table 2.[7]

The cointegration relationship is displayed in the first panel of table 2. The coefficients again have the expected signs. Specifically, the capital flow coefficient's negative sign indicates that foreign inflows into US treasury securities, measured as an increase in these securities' foreign holdings, exert downward pressure on long-term interest rates.

The second panel of table 2 shows the corresponding ECM. The adjusted coefficient of determination is 0.308 and, hence, hardly changes at all compared to the previous model. The negative and highly significant coefficient on the error-correction term confirms the cointegration relationship between the included variables. The coefficient on the first difference of foreign holdings again suggests a negative dynamic relationship between international capital flows and long-term interest rates.

[7] Foreign treasury holdings are used since the flow variable from the previous analysis is stationary and thus cannot be incorporated into the cointegration equation. However, note that the change in holdings used in the ECM represents a flow variable. Applying a Dickey-Fuller test reveals that foreign treasury holdings are I(1). A limitation of the approach is that the Johansen test is somewhat inconclusive regarding the number of cointegrating relationships between the variables.

Table 2: Estimation results – capital flows as equilibrium variable

Cointegration equation (dependent variable: $ltir_t$)							
$exin_t$	$excg_t$	rp_t	$flow_t$	*constant*			
1.335	0.520	-0.284	-0.126	2.533			
ECM (dependent variable: $ltir_t$): capital flows incorporated							
ect_{t-1}	$\Delta ltir_{t-1}$	$\Delta exin_t$	$\Delta excg_t$	$\Delta excg_{t-1}$	$\Delta excg_{t-2}$	$\Delta flow_t$	constant
-0.232	0.365	0.372	0.150	0.116	-0.119	-0.238	-0.014
***	***	**	**	*	*	**	
(-4.06)	(3.75)	(2.07)	(2.33)	(1.91)	(-1.96)	(-2.09)	(-0.33)

*,** and *** indicate significance on a 10%, 5%, and 1% level, respectively.

Since the above cointegration analysis suggests a long-run relationship between the considered variables, it is possible to estimate a "fair value" for bond yields by using the cointegrating regressions and comparing these fair values, as well as comparing them with the 10 year US treasury bond's actual yields.

Figure 4 shows the results of this exercise during the bond yield conundrum period. We can see that the fair value, which excludes international capital movements (dashed line), lies consistently above the actual 10 year yield (solid line), indicating an overvaluation of the bond market during this period. However, if international capital flows (dotted line) are included, this reduces the magnitude of the overvaluation during the conundrum period by an average of approximately 40 basis points, and also reduces the length of the overvaluation period considerably.

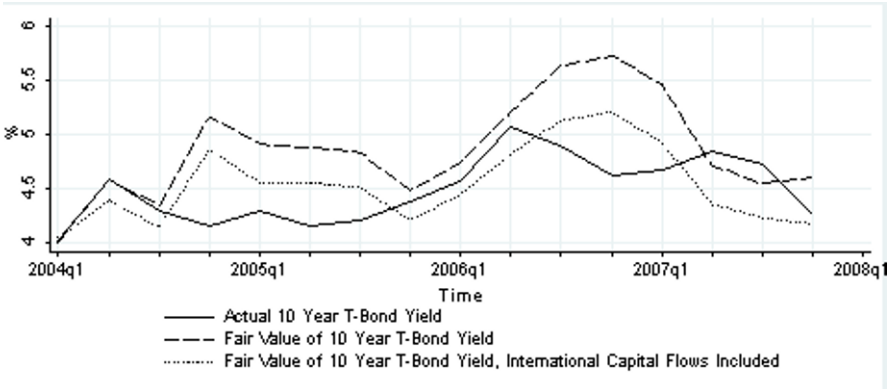


Figure 4: Actual ten-year treasury bond yield and fair values (source: own calculations with data from Federal Reserve Bank of St. Louis and Flow of Funds Account)

6 Global imbalances and the financial crisis

The above analysis shows that global imbalances, which materialized in the form of cross-border capital flows, contributed significantly to the low interest rate environment during the 2000s. Obstfeld and Rogoff (2009) emphasize that very low long-term interest rates were an important determinant of the recent financial crisis. This crisis certainly did not have a single cause but was the culmination of many contributing factors.[8] However, low long-term interest rates did promote the housing market bubble, which triggered the financial crisis when it burst.

According to Taylor (2009), the low interest rate environment already originated at the beginning of the 2000s when the Federal Reserve lowered short-term interest rates after the dotcom crash and kept them too low for too long. In line with the expectations hypothesis, long-term interest rates also declined (see Figure 1). This brought about low mortgage rates and rising house prices. Home owners borrowed increasingly against their housing equity and reduced savings, spurring a consumption boom, which led to seemingly robust growth rates in the US. But it also set a further widening of the US current account deficit in motion and, hence, capital imports (see Figure 2). The two developments therefore reinforced each other: Rising house prices stimulated borrowing from abroad, which in turn implied better access to foreign credit, as well as continuing low interest rates and, hence, further house price increases. Consequently, "[...] global imbalances did not cause the leverage and housing bubbles, but they were a critically important codeterminant" (Obstfeld and Rogoff, 2009, p. 1). The ensuing economic crisis, with deep recessions in a number of industrialized countries, suggests that global imbalances and a balanced growth path are conflicting phenomena.

As insinuated in the previous paragraphs, national policy choices play an important role in the increase in global imbalances. For example, policy measures to promote private saving in the US could have mitigated capital flows into the US economy. Likewise, a floating exchange rate for the Chinese currency – and, hence, less focus on export-led growth – could have dampened Chinese current account surpluses (Bini Smaghi, 2009). Policy coordination between countries is a crucial issue to control global imbalances and thus improve preconditions for a balanced growth path. However, getting many countries with different interests to agree as well as designing an institutional framework for discussing and committing to specific policy choices is hard to realize and beyond the scope of mere economic analysis.

7 Summary

This article dealt with global imbalances' influences on the development of long-term interest rates in recent years. In theory, arbitrage relationships between long-

[8] The Bank for International Settlements (2009) mentions, among others, management compensation schemes that favored short-term profits, regulatory weaknesses, and the mispricing of risk.

term and short-term interest rates (expectations hypothesis), inflation expectations, and real interest rates (Fisher equation) determine interest rates. Augmentations of these theories also consider risk aspects, leading to a risk premium as an additional determinant of interest rates.

Based on these insights, the empirical development of US long-term interest rates was examined. First, an overview was provided of the evolution of US long-term interest rates since the 1980s. This overview put the long rates' puzzling development between 2004 and 2007 – also called the bond yield conundrum – into context. The crux of this phenomenon was that long-term interest rates remained constant or even declined during a period in which the Federal Reserve raised short-term interest rates. This contrasted sharply with the Federal Reserve's past monetary tightening periods and was inconsistent with the pure expectations hypothesis, at least not if one assumes that economic agents actually correctly anticipated the central bank's series of interest rate increases.

Subsequently, the role that global imbalances play concerning long-term interest rates was examined. In the form of persistent and large international capital flows into the US, these could have driven down risk premia demanded by investors. A cointegration analysis was undertaken to estimate international capital flows' influence on long-term interest rates. The results suggested that international capital flows exert downward pressure on long-term interest rates both in the short and in the long run. Furthermore, international capital flows could explain a considerable part of the bond yield conundrum.

Global imbalances' influence on long-term interest rates contributed to the recent financial crisis, which originated in the US and spread across the globe. The resulting severe output losses indicate that global imbalances and balanced growth are incompatible. The coordination effort necessary to alleviate problems associated with global imbalances is largely a political matter.

References

Alexopoulou, I., Drudi, F., Scheithauer, J. (2006). What accounts for the low level of interest rates? *Background paper.* CGFS Working Group on Institutional Investors, Global Savings and Asset Allocation.

Allsopp, C., Glyn, A. (1999). The assessment: Real interest rates. *Oxford Review of Economic Policy*, 15(2), 1–16.

Bandholz, H., Clostermann, J., Seitz, F. (2009). Explaining the US bond yield conundrum. *Applied Financial Economics*, 19(7), 539–550.

Banerjee, A., Dolado, J., Galbraith, J.W., Hendry, D.F. (1993). *Co-integration, error-correction, and the econometric analysis of non-stationary data.* Oxford: Oxford University Press.

Bank of International Settlement (2009). *79th Annual Report: 1 April 2008– 31 March 2009.* Basel: BIS.

Bernanke, B. (2005). The global saving glut and the U.S. current account deficit. *Remarks at the Sandridge Lecture.* Virginia Association of Economics.

Bini Smaghi, L. (2009). The financial crisis and global imbalances – two sides of the same coin. *Speech at the Asia Europe economic forum conference*. Beijing.

Bofinger, P. (2001). *Monetary policy – goals, institutions, strategies and instruments*. Oxford: Oxford University Press.

Brière, M., Signori, O., Topeglo, K. (2006). Bond market 'conundrum': New factors explaining long-term interest rates? *Centre Emile Bernheim (CEB) Working Paper*, 06-024.

Caballero, R.J., Farhi, E., Gourinchas, P.-O. (2008). An equilibrium model of "global imbalances" and low interest rates. *American Economic Review*, 98(1), 358–393.

Case, K., Fair, M., Gärtner, M., Heather, K. (1999). *Economics*. Harlow: Prentice Hall Europe.

Clarida, R.H. (2007). Petrodollars, the savings bust, and the U.S. current account deficit. *Pimco Global Perspectives*.

Corsetti, G., Pesenti, P., Roubini, N. (1999). Paper tigers? A model of the Asian crisis. *European Economic Review*, 43(7), 1211–1236.

Crowder, W.J., Hoffman, D.L. (1996). The long-run relationship between nominal interest rates and inflation: The Fisher equation revisited. *Journal of Money, Credit and Banking*, 28(1), 102–118.

Dooley, M.P., Folkerts-Landau, D., Garber, P.M. (2005). Savings gluts and interest rates: The missing link to Europe. *NBER Working Paper*, 11520.

Engle, R.F., Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.

Evans, M.D., Wachtel, P. (1990). A modern look at asset pricing and short-term interest rates. *NBER Working Paper*, 3245.

Fisher, I. (1896). Appreciation and interest. *Publications of the American Economic Association*, 11, 1–98.

Frenkel, M., John, K.D. (2007). *Volkswirtschaftliche Gesamtrechnung*, Munich: Vahlen

Frey, L., Moëc, G. (2005). US long-term yields and forex interventions by foreign central banks. *Banque de France Bulletin Digest*, 137, 19–32.

Goodfriend, M. (1998). Using the term structure of interest rates for monetary policy. *Federal Reserve Bank of Richmond Economic Quarterly*, 84(3), 13–30.

Greenspan, A. (2005). Federal Reserve board's semiannual monetary policy report to the Congress before the committee on banking, housing, and urban affairs of the U.S. Senate. http://www.federalreserve.gov/boarddocs/hh/2005/february/testimony.htm

Gruber, J., Kamin, S. (2009). Do differences in financial development explain the global pattern of current account imbalances? *Review of International Economics*, 17(4), 667–688.

Hauner, D., Kumar, M.S. (2006). Fiscal policy and interest rates – How sustainable is the 'New Economy'? *IMF Working Paper*, 06/112.

Hicks, J.R. (1946). *Value and Capital.* Oxford: Clarendon Press.

Idier, J., Jardet, C., de Loubens, A. (2007). Determinants of long-term interest rates in the United States and the Euro area: A multivariate approach. *Banque de France Notes d'Etudes et de Recherche*, 170.

Ireland, P.N. (1996). Long-term interest rates and inflation: A Fisherian approach. *Federal Reserve Bank of Richmond Economic Quarterly*, 82(1), 21–36.

Krugman, P. (1998). What happened to Asia? *MIT Mimeo*.

Landau, B. (2000). Core inflation rates: A comparison of methods based on West German data. *Deutsche Bundesbank Discussion Paper*, 4/00.

Legg, A., Prasad, N., Robinson, T. (2007). Global imbalances and the global saving glut – a panel data assessment. *Reserve Bank of Australia Research Discussion Paper*, 2007-11.

Lucas, R.E.J. (1978). Asset prices in an exchange economy. *Econometrica*, 46(6), 1429–1445.

Obstfeld, M., Rogoff, K. (2009). Global imbalances and the financial crisis: Products of common causes. *Federal Reserve Bank of San Francisco Asia Economic Policy Conference Paper.*

Rudebusch, G.D., Swanson, E.T., Wu, T. (2006). The bond yield 'conundrum' from a macro-finance perspective. *Federal Reserve Bank of San Francisco Working Paper*, 2006-16.

Spahn, H.-P. (2006). *Geldpolitik – Finanzmärkte, neue Makroökonomie und zinspolitische Strategien*. Munich: Vahlen.

Stock, J.H. (1987). Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica*, 55(5), 1035–1056.

Taylor, J.B. (2009). *Getting off track*. Stanford: Hoover Institution Press.

Wolman, A.L. (2006). Bond price premiums. *Federal Reserve Bank of Richmond Economic Quarterly*, 92(4), 317–336.

Development and capital flow mysteries

Thomas Davoine

Abstract:

There are significant per capita income differences between countries. Since additional capital is more productive in firms and countries with low capital levels, standard theory predicts that poor countries should receive investments as a matter of priority. Nevertheless, capital flows from poor to rich countries, which constitutes a serious challenge to standard economic theory. Extensions to standard theory have been considered, such as considering countries' human capital qualities or financial development levels. Investors could, for instance, fear expropriation in poor countries and prefer lower but safer returns at home. The result is that both education and credit market imperfections help explain why capital does not flow from rich to poor countries, but theory still lacks something. Economic science must step up before it can make reliable economic development policy recommendations. Still, inaction would be worse, and incomplete scientific knowledge can be helpful in designing policy. While better financial institutions and education systems might both stimulate growth, these are not sure recipes for successful development. Policy-makers are invited into economists' kitchens.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_4,

Table of Contents

1 The capital flow mystery

Income differences are wide over time and between countries. The average per capita income in Switzerland in 2000 was 27 times the average in Nigeria. How these differences develop over time may be puzzling. While Nigeria and Thailand both had a per capita income about 11 times smaller than Switzerland in 1950, Thailand has closed the gap, while Nigeria has not: Switzerland's per capita income was 5 times that of Thailand in 2000, but increased to 27 times that of Nigeria. Why have some poor countries developed so much faster than others?

Economists have been working on this question for decades and have to date unfortunately not found any irrefutable answers. Of the various hypotheses put forward, many have been rejected, and none is universally accepted. In the quest for answers, the problem has been reduced to smaller pieces of which international capital flows is one such piece.

Neoclassical growth theory is the basic economic development theory and is now well accepted. It models production as a process that takes labor and capital as inputs and generates goods as outputs. These goods (or their financial equivalent, income) can either be consumed or saved. The resulting savings are invested to increase capital and, ultimately, output. For instance, a computer can either be used at home for leisure (consumption), or in an office to increase production.

Capital is much more mobile than labor: Investments can, in principle, be made anywhere, but there are cultural and political barriers to immigration. An interesting component of international economics is how capital moves between countries. Theory predicts that capital should flow from rich to poor countries. Yet, in the real world, the opposite is true. This seems to be a mystery. At first sight, the reason for capital not flowing from rich to poor countries appears less important than the question why some poor countries grow slower than others, but this is not so. If capital were moving from rich to poor countries, investment in poor countries would be higher and they would grow faster.

Why should capital flow from rich to poor countries? Consider the simple example of two relatively large farms identical in all but one respect: They own the same amount of land, produce the same type of crops, employ 20 workers apiece, but Farm *A* has four tractors, while Farm *B* has seven tractors. Both could really use eight tractors. Naturally, Farm *A* produces fewer crops per year. Suppose investors who are ready to invest the equivalent of one tractor come along. In which farm should they invest? In Farm *A*, because they can expect a larger increase in production than from Farm *B*, and therefore a higher return on investment. The same occurs at the country level. In theory, it is more profitable to invest in poor countries with little capital than in rich countries with ample capital. Therefore, we should see capital flowing from rich to poor countries, but we see the opposite. The observed direction of international capital flow is a mystery.

Owing to its importance for underdeveloped countries' economic growth, economists have been trying to explain this mystery. Education is a prominent hypothesis suggested by Nobel laureate Robert Lucas (see Lucas, 1990). In our example, it is possible that, at Farm *A*, only four of the 20 workers can operate a tractor, there-

fore the investors might be wise to invest in Farm *B*. Poor countries often not only have little capital, but also a weaker education system, which could explain why capital does not flow to them. While this hypothesis has been tested, the tests are only partially successful.

Several other hypotheses have been investigated, but to date none is entirely convincing. The financial sector's level of development is such a hypothesis. This hypothesis is presented in some detail to illustrate the distance between science and policy making, and why science may face difficulties when wishing to make clear-cut economic development policy recommendations.

2 Financial underdevelopment

A plausible reason for capital not flowing from rich to poor countries may be because investors from rich countries fear expropriation. Well-developed financial institutions, however, have mechanisms that protect investors. For instance, regulation can help investors recoup part of their assets if a company goes bankrupt or issues false financial statements.

The financial underdevelopment hypothesis is subsequently investigated in detail. The first goal is to assess whether this hypothesis can explain why capital does not flow from rich to poor countries. The second goal is to provide an illustration of what science can achieve and of the distance between science and policy making that has to be bridged.

This section seeks to meaningfully, yet accessibly, balance technical details, intuition, and overview. Readers wishing to avoid technical details can skim or skip Sections 2.2 and 2.3.

2.1 Overview of theory and results

The economic development process is generally modeled by neoclassical growth theory, a simple theory which assumes that output is the result of combining labor and capital as production factors. At a country level, the more capital, the higher the output and, hence, wealth. For instance, machines help workers produce more.

A prediction of this model is that investors should allocate funds where capital is scarcer because the returns will be higher. Using simple calculations, Lucas (1990) shows that the model implies that, in 1988, returns on investment were 58 times larger in India than in the US. However, in practice, capital does not flow from the US to India.

Investors may fear expropriation, or inefficient and costly work by financial intermediaries, so that foreign investors find it too risky or too costly to invest in poor countries. Theory has included the possibility of such credit market imperfections. The predicted capital flow direction can be overturned if credit market imperfections in a poor country are very high.

The following empirical questions arise: What is credit markets' de facto imperfection level? Is it large enough to reverse the predicted capital flow direction?

These questions are not easy to answer because credit market imperfection is impossible to measure directly from data. However, indices can be constructed to approximate the imperfection's strength.

The result is that the extended model with the measured parameters still predicts that interest rates are higher in poor countries. Although the model is still inconsistent with empirical evidence, it is nonetheless helpful. Without credit market imperfection, the simple model predicts that interest rates are on average 18 times higher in poor countries than in rich countries. When credit market imperfections are added, the model predicts that interest rates are nine times higher in poor countries. In line with empirical evidence, the model should predict that returns in poor countries are lower, since capital flows to rich countries. Taking credit market imperfections helps halve the gap: The ratio is not 18 to 1, but 9 to 1. This is a first step.

Note that models extended with the factor education also help us. The predicted ratio is no longer 18 to 1, nor 9 to 1, but 3 to 1. Together, education and credit market imperfections lead to an improved 1.5 to 1 ratio. It is very close to a 1 to 1 ratio, but still above it.

2.2 Theoretical model

Neoclassical growth theory is the basic theory of economic development for scientists. By means of a concise mathematical model, it can mimic the development of economies over time, taking a bird's-eye view. In particular, it leads to the conclusion that countries grow at a roughly constant pace once they have reached a stable equilibrium. This balanced growth property is consistent with empirical evidence.

The theory states that production is a function of two inputs, labor (L) and capital (K). The corresponding output (Y) can therefore be expressed as $Y=F(K,L)$. Different levels of productivity can also be considered, but this does not change the analysis's result. Since it approximates the de facto values well and is easy to handle, the production function (F) is given the so-called Cobb-Douglas form: $Y=F(K,L)=K^{\beta}L^{1-\beta}$, for a parameter value β between 0.3 and 0.4. Per worker variables simplify the analysis and are noted in lower case: $y=Y/L$ and $k=K/L$. Dividing both sides of $Y=K^{\beta}L^{1-\beta}$ by L gives $y=k^{\beta}$.

Investors who add 1 CHF to their capital can expect an average return equal to the additional output obtained. In mathematical terms, the additional output is the derivative of F with respect to K, also called the marginal product of capital. It provides the rate of return on investments (r). In this simple model, the interest rate is equal to the rate of return on investments. The calculation performed on $y=k^{\beta}$ delivers $r=\beta k^{\beta-1}$. Since $k=y^{1/\beta}$, the return can also be expressed in output per worker terms: $r=\beta y^{(\beta-1)/\beta}$.

Take two countries, one rich and one poor. Denote their corresponding variables with a R subscript for the rich country and a P subscript for the poor country. Then the ratio of returns is

$$\frac{r_P}{r_R}=\left(\frac{y_P}{y_R}\right)^{(\beta-1)/\beta}$$

This presentation follows that of Lucas (1990) – one of the most referenced articles in the economic literature. In this article, Lucas uses the US as the rich country example and India as the poor country example. He uses a value of 0.4 for β and notes that, in 1988, per capita income in the US was 15 times higher than in India, hence the ratio $y_P/y_R=1/15$. This means that the ratio of returns is $r_P/r_R=58$. In other words, the theory predicts that, in 1988, the return on investment in India was 58 times higher than the return in the US – a very powerful incentive for investing in India rather than the US! Nevertheless, capital has not been flowing from the US to India; the facts therefore contradict the theory.

A problem with this simple theory is that it neglects financial institutions' role. US investors might fear expropriation. Other aspects explaining this mystery could also be that Indian firms had problems accessing capital markets, or that Indian banks offered US investors a low interest rate for loans and Indian firms a high one.

To capture these reality elements, economists have extended the neoclassical growth model to allow for credit market imperfections. For instance, Matsuyama (2004, 2005) assumes that there are entrepreneurs who can invest a certain amount, for instance, *1*, in a business idea, but have insufficient personal wealth $w<1$. These entrepreneurs can either decide to borrow funds to pursue their business idea, or they can lend their wealth to the credit market and earn an interest at rate i. If they decide to borrow funds, the return on the investment will be the marginal product of capital, $r=\beta k^{\beta-1}$.

However, owing to credit market imperfections, investors will assume that only a fraction λ of the project revenue can be repaid on any loan. Since it is assumed that the investment needed was *1*, the project revenue is r X $1=r$, therefore the investors assume that only λr can be used to repay the loan of amount $1-w$. The investors can also obtain a return i if they lend on the market, therefore they will lend to the entrepreneurs if, and only if, the constraint $\lambda r \geq i(1-w)$ holds. In equilibrium, this constraint holds with equality.

When there are imperfections in the credit market, the interest rate i and returns on investments $r=\beta k^{\beta-1}$ are therefore related by the following relationship:

$$i = \frac{\lambda}{1-w} r$$

Why is this helpful? Again, take two countries, one rich and one poor. Assume that there are credit market imperfections in both countries, so that the interest rates and returns are related by $i=\lambda r/(1-w)$. The ratio of returns or marginal products of capital is – as before – equal to $r_P/r_R=(y_P/y_R)^{(\beta-1)/\beta}$. Thus the ratio of interest rates is equal to:

$$\frac{i_P}{i_R} = \frac{\lambda_P}{\lambda_R} \frac{1-w_R}{1-w_P} \left(\frac{y_P}{y_R} \right)^{(\beta-1)/\beta}$$

Suppose credit market imperfections are much more significant in the poor country, so that λ_P is a small number compared to λ_R. Then the ratio of interest rates will be low, possibly lower than 1, i.e. the rich country's interest rate is higher. Investors

would prefer to do business in this country and we would expect capital to flow to the rich country, as seen in the data.

The converse also holds: If credit market imperfections are comparable in both countries, the ratio of interest rates will be high and above 1. In this case, the poor country's interest rate is more attractive and we should see capital flowing to the poor country.

Whether credit market imperfections due to poor financial institutions really explain the mystery of why capital does not flow to poor countries is critically dependent on the various values of the parameter λ. Two empirical questions emerge: Is λ_P smaller than λ_R in reality? Is it sufficiently small so that $i_P/i_R<1$ for theory to predict that capital flows to rich countries, as it does in reality?

The next section provides early answers to these two questions.

2.3 Empirical analysis

In the theoretical model, the extent of credit market imperfections due to poor financial institutions is captured by a single parameter, λ, which represents the average percentage of profits that can be used with certainty to repay a loan. The lower the value, the larger the credit market imperfection, the less likely investors are to make loans to entrepreneurs. Are λ values really much lower in poor countries, thus explaining why capital flows from poor to rich countries?

Theory is consistent with the evidence of capital flowing from poor to rich countries if and only if

$$\frac{i_P}{i_R} = \frac{\lambda_P}{\lambda_R}\frac{1-w_R}{1-w_P}\left(\frac{y_P}{y_R}\right)^{(\beta-1)/\beta} < 1.$$

Estimates for credit market imperfection (λ_P, λ_R), per capita income (y_P, y_R), and entrepreneur wealth (w_P, w_R) will tell us whether the ratio i_P/i_R is indeed smaller than 1.

Some of these values are more easily observed than others. Per capita income (y) is easy to measure, whereas the two other variables are not. Not all countries have entrepreneur wealth measures. In the original Matsuyama (2004) model, w can represent either entrepreneurs' wealth or wage income. Since the size of investments is normalized to 1, the share of wages in output will represent the w measure.

The most challenging estimate is the index λ of credit market imperfections. While it is not directly observable, there are several ways to infer a value. This text presents one of these options[1] based on external financial dependence.

In a much-quoted paper on the relationship between financial development and growth, Rajan and Zingales (1998) demonstrate empirically that different industries have different external financing needs. Some industries, such as tobacco, generate high cash flows and do not need to rely on outside investors to finance new devel-

[1] Three other options to estimate the index λ of credit market imperfection deliver comparable results.

opments. Other industries, such as pharmaceuticals, have large investment needs and are dependent on external financing. Using industry data, they estimate for each industry (j) an *external financing dependence index* (EFD^j).

The mix and importance of industries differ between countries. To account for these differences, a weighted index is constructed for each country. A country's *EFD* will be the weighted average of the industry indices EFD^j, where the weight is the industry share (j) in national output[2].

Why would *EFD* be a good estimate of λ, and how good is it likely to be? The external financing dependence index is an indicator of how easy it is to obtain external financing. If a firm belongs to an industry that needs much external financing and it is difficult to obtain this due to credit market imperfections, it is less likely to survive. That industry's share of the national output will be lower, which will drag down the country index (*EFD*). In other words, large credit market imperfections will be reflected by a low *EFD* index. This is what is expected from an estimate of λ.

However, there are a number of assumptions that must hold for *EFD* to be a good indicator. For instance, one must assume that the skills of workers from the same industry are comparable in all countries. Rajan and Zingales (1998) discuss several of these assumptions to convince us that they should not invalidate the results.

As indicated in Table 2.2, in 1998, the OECD countries' *EFD* index was mostly close to 0.300, while it is often below 0.200 for developing countries.

As an illustration, take the respective values of the per capita income (y), credit market imperfection ($EFD=\lambda$), and wage share (w) of Bolivia and the US, and compare these with the theoretical predictions of models without or with credit market imperfections.

Table 2.1: Capital flow predictions between Bolivia and the US

		Bolivia	US	Relative
Data				
Per capita income (2003, $)	y	3,241	37,313	0.087
Entrepreneur resources (wage share)	w	0.36	0.58	0.621
Credit market imperfection (1998)	$EFD=\lambda$	0.144	0.346	0.416
Production parameter	β	0.4	0.4	-
Models' prediction of interest rate ratio				
No credit market imperfection		-	-	39.1
$r_P/r_R=(y_P/y_R)^{(\beta-1)/\beta}$				
With credit market imperfection		-	-	10.7
$i_P/i_R=(y_P/y_R)^{(\beta-1)/\beta}(\lambda_P/\lambda_R)(1-w_R)/(1-w_P)$				

[2] The following sources have been used for the numerical illustration: per capita income y is taken from the Penn World Tables (2006); the share of wages in output w is obtained from the UN National Accounts Statistics tables; the *EFD* calculations have been performed by Almeida and Wolfenzon (2005).

Table 2.1 shows two theoretical models' predicted interest rate ratio. The neoclassical growth theory model, where there is no credit market imperfection, predicts that the interest rate should be 39 times higher in Bolivia than in the US. The model with credit market imperfection predicts that the interest rate should be around 11 times higher. Both models predict that capital should flow from the US, a rich country, to Bolivia, a poorer country. The data show that capital flows in the opposite direction, so both models make wrong predictions. However, we have learned that models assuming that credit markets are perfect, actually do worse. Including credit market imperfections helps build models that are closer to reality, although there is still a gap.

Perhaps these results are specific to the country pair chosen for illustration. Table 2.2 below performs the same calculation with two groups of countries, averaging the values of each group. The first group includes less developed countries and the second group more developed ones[3].

Table 2.2: Capital flow predictions between less and more developed countries

		LDC	DC	Relative
Data				
Per capita income (2003, $)	y	4,071	28,411	0.143
Entrepreneur resources (wage share)	w	0.38	0.53	0.717
Credit market imperfection (1998)	$EFD=\lambda$	0.191	0.289	0.661
Production parameter	β	0.4	0.4	-
Models' prediction of interest rate ratio				
No credit market imperfection		-	-	18.4
$r_P/r_R=(y_P/y_R)^{(\beta-1)/\beta}$				
With credit market imperfection		-	-	9.2
$i_P/i_R=(y_P/y_R)^{(\beta-1)/\beta}(\lambda_P/\lambda_R)(1-w_R)/(1-w_P)$				

Table 2.2 leads to the same conclusion as table 2.1: Both models predict the direction of the capital flow incorrectly, as they predict that the interest rate in less developed countries is significantly higher. However, including credit market imperfections makes for more reliable predictions. In reality, we observe capital flowing from poor to rich countries, which the model would predict if the ratio of interest rates were below 1. Without credit market imperfections, the model predicts a ratio of 18; in the other case, it predicts a ratio of 9. Introducing credit market imperfection in the basic model therefore helps reduce the gap between theory and practice by factor 2.

[3] The group of less developed countries (*LDC)* comprises Bolivia, Colombia, Egypt, Ethiopia, Guatemala, Jordan, Kenya, Malawi, Mexico, Morocco, Nigeria, Tanzania, and Zimbabwe. The group of more developed countries (*DC)* is made of Austria, Belgium, Canada, France, Germany, Italy, the Netherlands, Spain, Sweden, the UK and the US. The choices were driven by data availability.

Similar results are obtained when credit market imperfections are incorporated into more sophisticated models. For instance, when the basic model takes education into account, the predicted interest rate ratio is 2.7 for the same group of countries. This is much closer to 1, but still larger than 1. When credit market imperfections are added to the model, the predicted ratio drops to 1.4. Again, a gap remains, but it is halved.

3 Three messages

What impact should theoretical and empirical results have on economic development policy? Can scientific knowledge help design policy that would help poor countries develop faster?

The previous section provided a glimpse into current scientific knowledge on economic development. It presented a hypothesis to explain why capital does not flow from rich to poor countries – weak financial institutions.

Putting this research into perspective, three messages emerge: first, financial institutions' development is a reason, but not the only one for capital not flowing from rich to poor countries; second, currently, economic science does not have an answer that explains *why* capital does not flow to poor countries nor why certain poor countries grow swiftly while other poor countries do not; third, because science still lacks reliable answers regarding economic development, policy-makers should not unreservedly accept scientists' advice on this topic.

3.1 Financial institutions' development matters, but not alone

Basic neoclassical growth theory predicts that returns on investment are significantly higher in poor countries, therefore capital should flow from rich to poor countries. Using 2003 values, it estimates that returns are 18 times higher in a group of less developed countries than in a group of more developed countries (see Section 2.3).

If the model is extended to allow for credit market imperfections due to weak financial institutions, the model still predicts that capital should flow to poor countries. The returns are estimated to be nine times higher.

These results lead to the following two interpretations: First, credit market imperfections help explain the capital flow direction because the gap is halved: Ideally, a model should predict a returns ratio below 1; the extended model halves the gap between 18 and 1. Second, credit market imperfections are not the only reason for capital not flowing to poor countries. Indeed, the model is still unable to close the gap.

Although credit market imperfections are not the only reason for slow economic development, policy can still be designed to help reduce these imperfections. Concretely, it depends on the specific forms of credit market imperfections.

Credit markets can be imperfect and retard investment for a number of reasons. For instance, the absence of a capital market; a lack of regulatory protection for

investors, who then fear expropriation; inefficient and costly capital allocation by financial intermediaries, such as banks; and corruption by some agents who appropriate resources for their own benefit.

There is no country in the world with perfect credit markets. If these markets were perfect, they would be so efficient that they would be costless. In this case, the interest rate banks would offer depositors would be the same as the that charged for loans to creditors. Clearly, this is not the case anywhere in the world.

Nevertheless, there are big differences between countries, which shows that there is room for improvement in some countries. According to the UN National Accounts Statistics tables, the average lending rate in Malawi was close to 50% in 2003, compared to an average of 4% in the US.

Ultimately, the concrete measures to reduce credit market imperfection will depend on its various causes. These can be a public *and* a private matter. For instance, investor protection measures are often a question of regulation. These can consist of requiring tighter financial disclosures from creditors. On the other hand, to the extent that private agents own such intermediaries, financial intermediaries' capital allocation is primarily a private matter.

Not all of the possible financial development measures are listed here. The vast literature on financial development and growth's effect should be helpful in this regard. The theoretical and empirical analyses show that the financial development level matters, but that no one should expect a miracle from the strengthening of financial institutions, which is by far not the only factor influencing foreign investment in developing countries. We have seen that education matters, too. The next section will present other possible factors.

3.2 Science does not (yet) have all the answers

The financial development level matters for economic growth, but it is not the only factor. As noted, education level also matters. The empirical analysis shows that the predicted interest rate ratio could be reduced from 18 to 1.5, but there is still a gap, with level 1 being consistent with the observed capital flow direction between poor and rich countries. However, the neoclassical growth theory model must be enriched with more than education and credit market imperfections to be consistent with the data. What else could explain the reason for capital not flowing from rich to poor countries?

Many other factors have been considered to explain per capita income differences between countries, and these can also explain the capital flow direction. For instance, there could be differences in productivity, in the intensity of the education, in life expectation, in health, and in capital depreciation.

Caselli (2005) presents a fairly recent review of the economic literature that tries to explain cross-country income differences. The conclusion is that we know some of the main suspects, but not all of them. As the author states, "a sentence commonly used to summarize the existing literature sounds something like: differences in efficiency account for at least 50% of differences in per capita income" (p. 681)

where efficiency is an indication or measure of how much output the production process obtains from inputs, such as labor or capital.

One challenge for today's economists is to be more specific about what efficiency is. Other terms for efficiency are productivity and technological progress. Abramovitz (1956) was the first to call efficiency a "measure of our ignorance," because it measures what is left to explain once differences in production inputs have been considered. Many authors – including Caselli (2005) – are currently sympathetic to this viewpoint. Technology, productivity, and efficiency are all related concepts that are difficult to define and measure precisely.

Nevertheless, progress has been made; it is hoped that more will follow. The quality of education seems to make little difference, but the quantity does. There is no point in comparing efficiency at an aggregate, national level, and making efficiency comparisons between different countries at the industry level is preferable. The example contrasting Nigeria and Thailand (both equally poor in 1950, but Thailand was five times richer in 2000) is instructive. Science tells us that a reason for this difference is education. UN data show that in 2005 adults in Nigeria had an average of five years of schooling, while those in Thailand had an average of six. While this difference in education is consistent with the theory, it is unlikely to explain why Thailand developed much faster.

This discussion shows that science may provide some answers regarding why some poor countries develop and why some do not, but it by no means has all the answers, and much remains to be explained. This conclusion regarding cross-country income differences echoes the one regarding capital flow directions.

Consequently, there is a gap between the current body of scientific knowledge and what would be desirable for reliable policy recommendations. Is this a reason for inaction or a lack of policy intervention? No. While scientific knowledge is incomplete, it can still be helpful. On the whole, financial development and education matter for economic development. Nevertheless, it should not be forgotten that this is incomplete knowledge. Today, policy recommendation regarding economic development is as much an art as a science. The development of financial institutions, including capital market extensions, investor protection, anti-corruption measures, and more efficient banking intermediation, should be supported. But this alone will not help poor countries develop. While it can take decades for education to have a significant effect, it also matters. Other factors, which science cannot as yet identify precisely, certainly matter too.

3.3 Take a look inside the kitchen

Policy-makers should not take expert advice on poor countries' economic development at face value. As noted in the previous sections, too many unanswered questions remain. Instead, policy-makers should lift scientists' veil of expertise, and should make their own assessment of economists' hypotheses, data, and analyses. It may, for example, be acceptable to take a physicist's advice on satellite trajectories without lifting the veil of expertise but this does not hold for an economist's advice on economic development.

There is a large degree of consensus among physicists and other natural sciences, even if there are debates on open scientific questions. The degree of consensus is much lower in economic science, especially on the topic of poor countries' economic development. You can eat a meal prepared by a physicist without worrying: Unpleasant surprises are unlikely, whoever the physicist is who prepares the meal; there is no need to look inside his kitchen. In contrast, policy-makers should always take a look inside a development economist's kitchen: The quality of the meal is not guaranteed, as there are many unknowns in the process; two different economists may serve you the same meal but their taste will differ completely. If the process is problematic, the policy-maker should not hesitate to consider different cooks, kitchens, and meals.

Is this feasible? Will a kitchen visit put a policy-maker in a better position? Yes, under certain conditions. The presentation of financial institutions' role in economic development in Section 2 serves as an illustration: Readers with a university education in social sciences, natural sciences, or engineering should have few problems understanding the analysis.

However, there are certain conditions to lifting the veil of expertise or looking inside the kitchen. The language of science differs from everyday language and is not always accessible. The presentation in Section 2 involved a simplification of some complicated scientific papers. Certain assumptions were glossed over to keep the language understandable.

If a policy-maker relies on a scientist to translate scientific knowledge into everyday language, the benefit of lifting the veil of expertise may be seriously compromised for at least two reasons. First, some scientists' personal agendas could color the translation, but succeed in hiding this bias, so that the translation passes as an unbiased scientific recommendation. Second, there are many topics on which the degree of consensus is low, or the number of unanswered questions is large; consequently, where two economists could provide two very different stories. Policy-makers and their staff should have sufficient exposure to scientific language to grasp economists' knowledge and recommendations without translation, and should make their own assessments.

Policy-makers are presumably too busy to maintain proficiency in scientific knowledge, and many may rely on staff or external support in this regard. This support comes with conditions, however: Neither the staff nor the external support should be involved in the production of scientific knowledge. This will maintain policy-makers' independence and guarantee transparent assessment of the different alternatives available to support poor countries' economic development.

References

Abramovitz, M. (1956). Resource and Output Trends in the United States Since 1870. *NBER Chapters*, National Bureau of Economic Research, Inc, 1-23.

Almeida, H. & Wolfenzon, D. (2005). The effect of external finance on the equilibrium allocation of capital. *Journal of Financial Economics*, 75(1), 133-164.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In P. Aghion & S. Durlauf (Ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 9, (pp. 679-741). Elsevier.

Heston, A., Summers, R. & Aten, B. (2006), Penn World Table Version 6.2, CIC, University of Pennsylvania. http://pwt.econ.upenn.edu/

Lucas, R.E. (1990). Why doesn't capital flow from rich to poor countries? *American Economic Review*, 80(2), 92-96.

Matsuyama, K. (2004). Financial Market Globalization, Symmetry-Breaking, and Endogenous Inequality of Nations. *Econometrica*, 72(3), 853-884.

Matsuyama, K. (2005). Credit Market Imperfections and Patterns of International Trade and Capital Flows. *Journal of the European Economic Association*, 3(2-3), 714-723.

Rajan, R.G., & Zingales, L. (1998). Financial Dependence and Growth. *American Economic Review*, 88(3), 559-586.

Sustainable development of housing finance markets – An international perspective after the crisis

Friedemann Roy, The World Bank Washington DC[1]

Abstract:

Like Michel Ende's novel "The Never-ending Story," it seems that house price bubbles tend to occur at regular intervals. Due to the interlinkages with the banking sector, a bursting property bubble often leads to a banking crisis, damping economic growth and destroying prosperity. The objective is to show the relationship between property and bank lending. The factors which fuel a housing bubble will be discussed and the indicators pointed out that help identify rising house price inflation. In conclusion an analysis is given of the future prospects of property markets in advanced economies and in emerging markets.

[1] The findings, interpretations, statements, and conclusions expressed herein are those of the author alone and do not necessarily reflect the views of the International Bank for Reconstruction and Development/The World Bank and its affiliated organizations, or those of the Executive Directors of The World Bank.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_5,

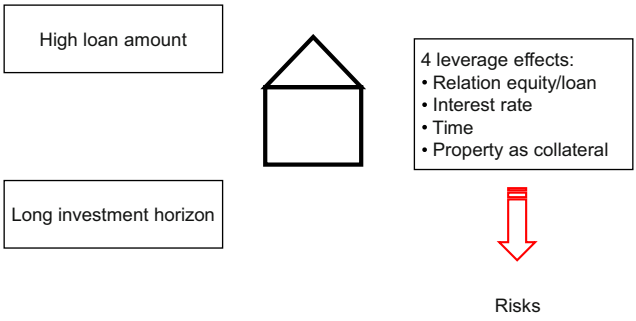
Contents

"Banks are leveraged and property is leveraged, so there is double leverage," says Brian Robertson of HSBC.[2] As will be shown, property cycles are closely related to banking activities. The previous six banking crises in advanced economies before the latest crisis in 2008, had property at their heart (Spain in the 1970s, Norway in the 1980s, and Sweden, Finland, the UK, and Japan in the 1990s). These crises tended to happen at the peak of housing booms, or just after a price collapse.

Subsequently, the specific characteristic of housing investments financed with mortgage loans will be outlined; thereafter the sources of housing bubbles will be discussed, followed by indicators that could identify a housing bubble. Lastly, the prospects for the future development of property markets will be discussed, including recommendations on how future bubbles could be prevented. Examples from advanced economies and emerging markets will be used throughout.

1 A house – a risky and volatile asset

A house is typically the biggest single asset of a private individual. Typically, 80-90% of the purchase price or construction cost is financed by a loan. This financing structure exposes the borrower to a number of risks. Chart 1 serves as an illustration:



Source: Roy

Chart 1: Leverage effects in a housing finance transaction

- High leverage and fluctuating interest rates. Only a thin layer of equity is available to protect the borrower. In case house prices fall, the borrower will be quickly pushed into negative equity. The term structure of the mortgage loan may reinforce this effect, especially if interest rates are variable or can only be fixed for shorter duration than the agreed financing term. If interest rates increase, the borrower may not be capable of repaying the mortgage loan.
- Duration of the loan. The duration of a mortgage loan is typically 25 to 30 years. This time span may involve a set of unknowns which could lead to

[2] Bricks and Slaughter – A special report on property. *The Economist*, March 5, 2011.

a borrower's default (e.g. unemployment, sudden increase in interest rates, etc.).
- A dwelling or a house is an indivisible good. In comparison with a share portfolio which allows an investor to sell parts of his or her holdings, a borrower cannot sell a part of his or her house (e.g. kitchen) in case of payment difficulties.

The leverage at the individual household level is reflected in the economy. The aggregate value of property held by private households in developed countries amounted to about 48 trillion USD in 2002 and it reached its peak at around 60 trillion USD in 2007. After the so-called collapse of the global financial system in 2008, the value was still estimated at about 52 trillion USD, which accounts for about 67% of the developed countries' combined GDP.

Most of this wealth can be found in the banks' balance sheets. Banks are especially prone to fluctuations in the property markets because property is regularly offered as collateral in a loan agreement. Not only private households offer property as collateral but also small and medium-sized companies tender property as collateral in loan applications. As a result, property accounts for a major portion of the financial system's aggregate balance sheet.

Another reason for the considerable share of property in the banks' balance sheets is that mortgage loans are subject to lower capital charges than most other assets. This regulation is based on the assumption that the property will retain some value if the borrower defaults. The apparent safety of property as an asset class may lead to the false conclusion that lending against property is prudent. Although collateralized lending offers a certain protection to the individual lender, it implies considerable systemic effects:

Property reinforces pro-cyclical effects in the economy. In a boom, rising property prices increase the value of the collateral held by banks, which usually incentivizes banks to increase their lending activities. The availability of easier credit means that property can sell for more, driving up house prices further. This correlation works in the opposite direction as well. When house prices fall, lenders react with tighter underwriting standards, forcing borrowers to sell and accelerating the decline in prices. Given the importance of property in the banks' balance sheets, losses from real estate busts are likely to be synchronized across banks.

Property absorbs an oversupply of credit. The safety net governments have placed under financial systems after successive crises has permitted banks to borrow and lend more against a shrinking capital base in pursuit of higher returns. A great deal of the free liquidity was absorbed by lending backed by property as collateral. Assessments from rating agencies gave additional reassurance that risks in property were low.

Besides leverage, the inefficiency of property markets is another reason for property being a risky and volatile asset. Property appraisals are subjective, and the price of a property is set locally by recent transactions. The value of any particular home,

and the amount that can be burrowed against it, is largely determined by whatever a similar house nearby sells for. In addition, houses are not fungible goods like equities or bonds. A uniform price for a number of houses cannot be determined. The law of one price does not apply to properties.

The absence of fungibility does not allow for short sales.[3] Individual properties and neighborhoods differ, which make it difficult to construct accurate hedges. For example, it is impossible to borrow the Empire State Building in order to sell New York real estate short.[4] Short sales are an important indicator, however, in identifying overpriced goods. Short sales were behind the identification of the ENRON scandal.[5]

In conclusion, systemic effects of property prices reinforce pro-cyclicality within an economy and have a considerable impact on the banking sector, households and the government, mainly caused by the leverage in lending with property as collateral.

The pro-cyclical behavior is the result of an absence of short-sales and disaster myopia which is defined as the economist's argot for the tendency to underestimate the problem of low frequency shocks which often occur at the beginning of an upcoming housing bubble. Another factor is the interaction of myopia and competition. When lending by myopic banks pushes property prices above fundamental values, more prudent banks are either priced out of the market or forced to drop lending standards. Other myopic banks see the high returns and enter the market. As the returns of the banks start to erode, banks take on more leverage and risk to protect their return on equity. Disaster myopia then leads to disaster magnification in the downturn.[6]

A housing policy framework which supports homeownership could be another factor which incentivizes more investments in property. This is described in more depth in the next section.

2 Sources of housing bubbles

"Rekindling the dream of homeownership for America's working families can prepare our nation to embrace the rich possibilities of the twenty-first century," as Pres-

[3] A short sale is the sale of a stock someone does not own. Investors who sell short believe the price of the stock will fall. If the price drops, you can buy the stock at the lower price and make a profit. If the price of the stock rises and you buy it back later at the higher price, you will incur a loss (definition from U.S. Security and Exchange Commission, http://www.sec.gov/answers/shortsale.htm).

[4] Bricks and Slaughter – A special report on property. *The Economist*, March 5, 2011.

[5] The Enron scandal, revealed in October 2001, eventually led to the bankruptcy of the Enron Corporation, an American energy company based in Houston, Texas, and the dissolution of Arthur Andersen, which was one of the five largest audit and accountancy partnerships in the world. In addition to being the largest bankruptcy reorganization in American history at that time, Enron was attributed as the biggest audit failure (from Wikipedia, the free encyclopedia).

[6] Plender, J. Overarching Problems. *Financial Times*, January 27, 2011.

ident Clinton wrote in the preamble to the National Home Ownership Strategy, the goal of which was to expand home-ownership (in 1995). As evidence mounted in the early 1990s that more and more Americans faced stagnant or declining incomes, the political establishment started looking for ways to help them with fast-acting measures. Unequal access to quality education was identified as one major reason for the rising income inequality. The provision of affordable housing, especially for low income groups was considered an obvious answer to this problem.

With the help of Fannie Mae, Freddie Mac, the Federal Housing Administration (FHA) and the Department of Housing and Urban Development, the access to cheaper credit could be realized, setting the stage for a boom in low-income housing construction and lending. In June 2008, Fannie Mae and Freddie Mac, FHA and other government programs were exposed to about 2.7 trillion USD in loans,[7] approximately 59% of total subprime and Alt-A loans.[8] Initially, this housing policy was a success. The homeowner ship rate increased to nearly 70% in 2004.[9]

A monetary policy concentrating on low interest rates provided further support to the favorable housing policy framework.[10] Additionally, the development of new financial instruments which allowed many new borrowers to enter the mortgage market, as well as more enhanced and more sophisticated securitization techniques incentivized investors to take on higher credit and liquidity risks. There was a widespread belief that financial engineering was capable of taming risk by better tailoring exposures to investors' risk preferences. The securitization of subprime mortgages combined with a heavy appetite for these instruments in countries such as Germany, Japan, China and other emerging markets fuelled the perception that house prices would continue to climb forever.

A couple of other factors undermined the perception that the rise in house prices in the U.S. was on a stable footing:[11]

- The United States was considered special because of the world's most reliable system of financial regulation, the most innovative financial system, a strong political system, and the world's largest and most liquid capital markets. It could withstand huge capital inflows without worry.

---

[7] An Alt-A mortgage, short for Alternative A-paper, is a type of US mortgage that, for various reasons, is considered riskier than A-paper, or "prime," and less risky than "subprime," the riskiest category. Alt-A interest rates, which are determined by credit risk, therefore tend to be between those of prime and subprime home loans. Typically Alt-A mortgages are characterized by borrowers with less than full documentation, lower credit scores, higher loan-to-value ratios, and more investment properties.

[8] Rajan, R. G. (2010). *Fault Lines – How hidden fractures still threaten the world economy*. Princeton and Oxford: Princeton University Press, p. 38.

[9] By 2010 it decreased by 5 percentage points to 65% as result of the collapse of the housing markets (according to latest Census of April 2010).

[10] For a more in-depth discussion on US interest rate policy, see the chapter written by Daniel Kienzler.

[11] Reinhart C. M., & Rogoff K. S. (2009). *This Time is Different – Eight Centuries of Financial Folly*. Princeton and Oxford: Princeton University Press, p. 214.

- In addition to these strengths, the United States was seen as having superior monetary policy institutions and monetary policy makers. One fact which was overlooked was that most of the mortgages were transferred into trusts which were not supervised and regulated by the banking supervisors like the Federal Reserve ("shadow banking system").
- Because of increased global financial integration and deeper global capital markets, it was possible for countries to go deeper into debt. In a time of low interest rates, securities with mortgages as underlying collateral were considered an alternative to government bonds offering a higher return by only minimally increasing risk.
- Given all these strengths, the capital markets of the United States offered to rapidly emerging economies a secure place to invest their surplus funds for diversification purposes.

The availability of cheaper credit did not only fuel a housing boom in the United States, but also in Western, Central and Eastern Europe. For example in Serbia, the volume of mortgage loans grew by 244% in 2005. In Bulgaria, the ratio of mortgage lending to GDP rose by 1.4% (2004) to 2.1% (2005), while house prices surged by 47% and 36.5% in the two countries during the same years.[12] Chart 2 illustrates the trend in Bulgaria.

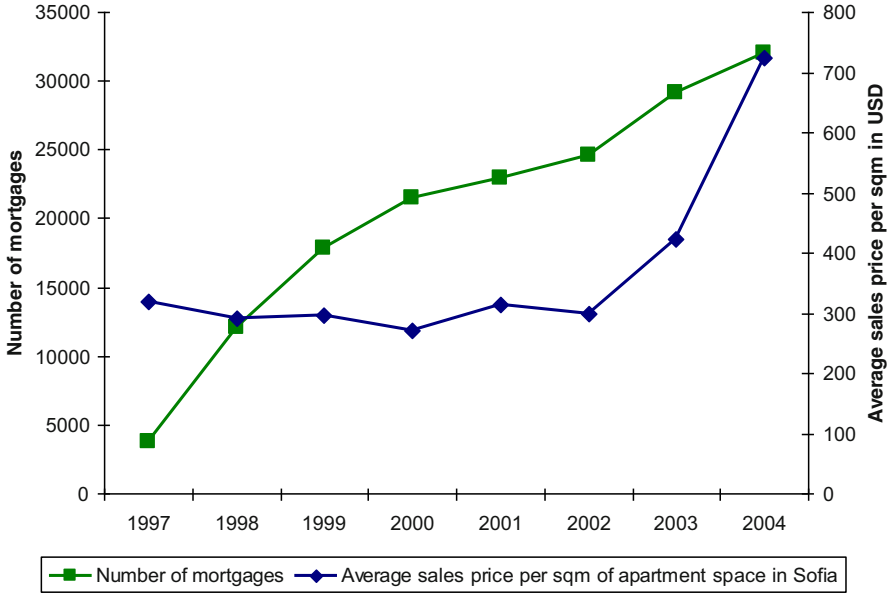


Chart 2: Correlation between supply of mortgages and house prices in Bulgaria (1997–2004)

[12] EMF Hypostat 2005 (2006). A review of Europe's Mortgage and Housing Markets. European Mortgage Federation: Brussels, p. 129 and p. 138.

Macroeconomic reforms in Central Eastern Europe helped stabilize interest rates. Another important factor was the availability of cheap loans denominated in foreign currency. These loans triggered a lending boom, especially in Hungary, Romania, Bulgaria, and the Baltic States. Chart 3 shows the increased share of mortgage loans denominated in foreign currency compared with all mortgage loans. A third factor was a loosening of underwriting standards to attract more borrowers.

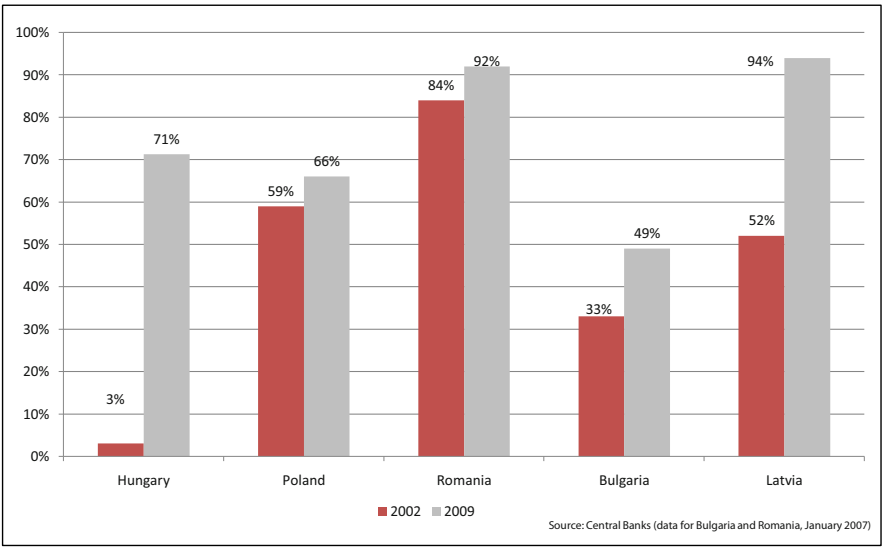


Chart 3: Share of mortgage loans in foreign currency to overall mortgage loan portfolios

In Hungary, for example a considerable share of the loan portfolio was denominated in foreign currency: about 66% of all loans were granted in Swiss francs (CHF) and, consequently, borrowers were exposed to significant foreign currency risk. The rapid expansion in mortgage lending was mainly fuelled by two elements:

- Generous subsidies (from 2000 to 2004). In 2000, two interest rate subsidies were introduced:[13] (i) an interest rate subsidy on mortgage bonds and (ii) an interest rate subsidy for loans which were connected with new construction. The third element in the subsidy program was the personal income tax mortgage payment allowance.[14]

[13] See, Hegedus, J. (2009). *Housing Policy and the economic crisis – case of Hungary, Budapest*, p. 7.

[14] This subsidy entitled the borrower to deduct 40% of the mortgage repayments (principal and interest) up to a maximum of 240,000 HUF (about 1,300 USD) from the tax payment. See Hegedus, J., & Somogyi, E. (2005). An Evaluation of the Hungarian Mortgage Program 2000–2004. In Struyk, R., & Hegedus, J. (Ed.), *Housing Finance – New and Old Models in Central Europe, Russia and Kazakhst*an. Open Society Institute: Budapest, p. 192.

- Cheap funding for banks in foreign currency (2004 to 2008). Since the cost of the mortgage subsidy program exceeded estimations,[15] the government cut back the subsidies. However, this did not hold back the further expansion of the market due to the access of Hungarian banks to cheap funding in foreign currency (fx).

With the boom accelerating in 2008, most lenders lowered mortgage lending standards. Lenders were not able to compete on interest rate, as the aggressive offer of fx loans had already substantially lowered interest rates on mortgages and further reductions would have eliminated their lending margins. In particular, loan-to-value ratios[16] and payment-to-income ratios were increased. It was also regular practice to take into consideration informal incomes in the creditworthiness assessment of the applicant. Another feature of mortgage loan origination was the widespread use of brokers by most lenders to attract borrowers. Although brokers did not grant the loans, the loan quality of loans originated by them was considered lower than those originated in the bank branches.

One reason for the attractiveness of borrowing in foreign currency was the interest rate difference between loans denominated in Swiss francs (CHF) and Hungarian forints (HUF) although borrowers did not take into consideration the currency risk. For a mortgage loan denominated in CHF, the interest rate was in the range of 4-5%, whereas for HUF denominated mortgage loans, the interest rate was about 16%.

In conclusion, a housing policy framework which focuses only on the support of homeownership and ignores economic realities is likely to fuel a housing bubble. Housing policy makers are advised to balance carefully their policy goals against the possibility of creating a bubble. Policies which limit, for example, loan-to-value ratios and affordability ratios may help to achieve this balance. Other supportive factors are loosening of underwriting standards and the availability of cheap loans.

Interestingly, the magnitudes of the declines in real housing prices from peak to trough are not appreciably different in emerging market and advanced economies. Which indicators can be used to identify and explain a housing bubble? The next section aims to provide clarity to this question.

3 Indicators to identify housing bubbles

When the housing bubble in the United States burst, the repercussions were felt throughout the whole world. When the investment bank Lehman Brothers declared bankruptcy in September 2008 and considering the following events, it became

[15] By 2003, the Government spent on subsidies about 153 billion HUF (against a projected cost of 42 billion HUF by 2005).

[16] The loan to value or LTV ratio of a property is the percentage of the property's value that is mortgaged. If one divides the mortgage amount by the lesser of either the appraised value or the selling price, one gets the LTV.

clear that the collapse of the U.S. housing market had been transformed into a global financial crisis.

In Western Europe, especially Spain, Ireland and the UK, also experienced a collapse of their housing markets. Countries like Germany, Austria and Sweden imported the crisis, either through investments of their banks in U.S. mortgage backed securities or through their exposure to bank lending in Central and Eastern Europe where the markets also went bust. Overall, the fall in house prices erased more than 3.4 trillion USD of household wealth in 2008.[17] Chart 4 shows the development of house prices in Western Europe before and after the crisis.

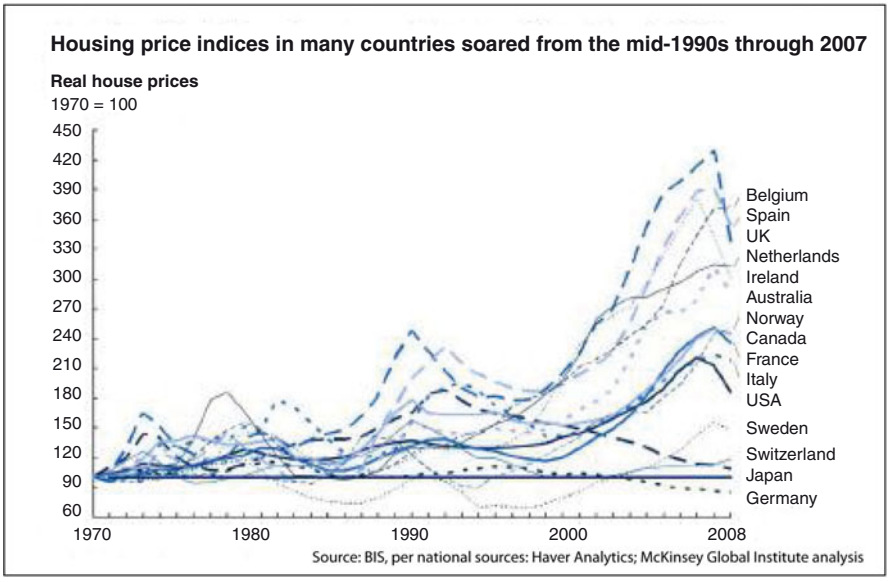


Chart 4: Development of house prices in Western Europe from 1970 to end 2008

In emerging markets, a different picture emerged in the aftermath of the global financial crisis. As shown in Table 1, three scenarios can be identified:

Table 1: Different developments in emerging markets after the crisis

Downward trend	Stable development	Upward development
Hungary, Romania	Most countries in Sub-Saharan Africa	Brazil
Mexico	Korea, Malaysia	Jordan, Lebanon
South Africa	Colombia	China

Source: Roy

[17] Roxburgh, C., Lund, S., et al. (2009). *Global capital markets: Entering a new era*. McKinsey Global Institute: Washington, Seoul, San Francisco, London, p. 11.

1. Downward trend. As already discussed above, many markets in Central and Eastern Europe contracted during the global financial crisis. For example, house prices in Estonia fell by about 20%. Similarly hard hit were Mexico and South Africa. One of the main reasons was that long-term funding for mortgages was no longer available. South Africa experienced considerable capital outflows and in Mexico, private lenders exited the market.
2. Stable development. The global financial crisis had only a limited impact on housing markets. Either these countries were not integrated in the global capital markets (most countries in Sub-Saharan Africa, except Nigeria, Kenya and South Africa) or these countries had already experienced an economic contraction in earlier years which had been caused by the collapse of their domestic housing markets, and they had implemented appropriate measures aimed at avoiding these crises (e.g. Thailand, Colombia).
3. Upward development. In comparison with other countries, these countries are experiencing economic growth followed by rising incomes and increased availability of mortgage loans (e.g. Brazil, China). Another reason is the gap between supply and demand for houses which has pushed house prices upward (e.g. Lebanon).

House prices in China have risen over 53% over the last four years. In 2009, they increased by 25% and in 2010 an average increase of 19% was recorded. The rise was driven by a rapid expansion of liquidity (due to a fiscal stimulus package and monetary expansion) as stock market prices were less attractive and the real return on bank deposits was negative. Although rising incomes and an ongoing urbanization would justify rising prices, an assessment whether the recent price rises indicate a housing bubble, appears warranted.

The following indicators serve to identify a housing bubble:

Development of price-rent ratio: The development of the price-rent ratio is of particular interest for investors because a rising ratio suggests that an increasing share of an investor's expected return stems from expected capital appreciation, i.e. from potential future rental growth, rather than from current rental returns. This implies that speculative motives play an increasing role. According to Deutsche Bank Research, this indicator points to a sizeable overvaluation of house prices for most regions in China.[18]

House price income ratio: This measures the affordability of housing. Data sets indicate that prices have grown much faster than income, i.e. overall housing affordability has declined. Deutsche Bank Research calculated that it would take

[18] Clemens, U., Dyck, S., & Just, T. (2011). *China's housing markets: Regulatory interventions mitigate risk of severe bust*. Deutsche Bank Research: Frankfurt, p. 3.

almost 20 years of annual income growth of 10% for China to reach similar affordability ratios as Germany or Spain.

Supply side developments: The number of houses completed and vacancy rates are suitable indicators to describe changes in the supply of houses. House price rises and the increased availability of credit is poised to drive supply of housing. The Chinese Government has recently implemented a number of measures (e.g. introduction of a property tax in the first quarter of 2011, home purchase bans, mortgage restrictions, etc.) to curb speculative demand (especially in the luxury segment), but has also tried to stimulate the supply of public and affordable housing. In 2010, the overall supply of housing decreased.

Rising vacancy rates indicate a growing oversupply of houses. Currently, overall vacancy levels are low in all Chinese cities. The highest vacancy rate amounts to 4.5 % of the housing stock (in Yinchuan).[19] Thus, the supply side indicators do not signal an overvaluation.

Demographic trends: Population growth and the urbanization rate serve to describe the impact of demographic trends on house prices. A strong population growth is an appropriate indicator of rising house prices. China's population is about 1,3 billion and is anticipated to peak at 1,4 billion by 2030 and then slowly start declining.[20]

High urbanization rates signal an increasing demand for housing in cities. According to recent estimates,[21] the urban population in China is expected to increase from 572 million in 2005 to 926 million in 2025. By this date, 221 cities are expected to have more than 1 million inhabitants each and 8 megacities with a population exceeding 10 million each. The massive urbanization will put substantial pressure on availability of land, financing and the development and construction industry. Pressure on house prices is expected to come from the urbanization trend. Population growth may not play an important role.

Availability of credit: The increased availability of housing credit may lead to higher house prices. From 2004 to 2009, the volume of outstanding mortgage loans increased from 1,500 billion RMB to 4,600 billion RMB (about 710 billion USD). The ratio of mortgage loans to GDP amounted to about 15.3% (in 2010).[22] Mortgage loans and loans to developers accounted for 20% of total outstanding loans. Lending standards are understood to be prudent. However, these data may conceal a less positive picture:

[19] Clemens U., Dyck S., & Just T. (2011). *China's housing markets: Regulatory interventions mitigate risk of severe bust*. Deutsche Bank Research: Frankfurt, p. 6.

[20] Rosenberg, M. (2010). China's population. http://geography.about.com/od/populationgeography/a/chinapopulation.htm.

[21] Renaud, B. (2009). *A Systemic View of Housing in China's New Urban Area*. Seoul: Presentation at KDI School.

[22] Home Truths: China's property market. *The Economist*, May 27, 2011.

- It is probable that property appraisal values in loan agreements have been overstated to allow for higher loan-to-value ratios and hence higher loan amounts.
- A sharp increase of wealth management products has fueled securitization. Banks move loans off-balance sheet and obtain capital relief by packaging them into wealth management products sold to investors through trust companies. This form of securitization is capable of hiding the true performance of loan portfolios and may distort data on defaulted loans.
- The real exposure to public sector debt through "local government financing platforms" is difficult to determine. Not being allowed to borrow directly, many local governments circumvent the prohibition by establishing companies to borrow on their behalf, using land as collateral. If there is a sudden decline in land prices, the banks which lend to these companies could see their capital base eroding due to an increasing number of defaults from this source.[23]

In conclusion, a number of indicators help identify a bubble. For China, a clear answer cannot be given due to wide differences among cities. House prices in the top 5 cities (Shanghai, Beijing, Guangzhou, Shenzhen, Tianjin) have risen above fundamental justified levels, leading to decreased housing affordability and implying political and social risks.

It seems that government policies have been effective in slowing overall house price growth in the country. However, the risk of a severe bursting bubble is still apparent. Even in case of a more benign slowdown, the construction sector and related industries are facing a potentially big negative impact. Additionally, the effect on the banking sector is difficult to determine as the level of distorted property valuations of those properties offered as collateral is unknown and the debt levels of local governments are not disclosed.

It is considered unlikely that the situation in China will develop along the lines seen in the United States and Spain. Pent-up demand and overall macro-economic dynamics are still strong. Although the exact exposure of the banking sector is impossible to determine, it is at least limited to Chinese lenders. Repercussions are therefore unlikely to be felt by banks in other countries.

China is presented as an example of a booming housing market. Another booming market is Brazil. Their current experience is the opposite of what is occurring in most developed countries. When will these markets in these countries return to their growth paths again? What are the prospects of housing markets? The next section provides interesting answers.

[23] Renaud, B. (2011*). Housing Bubbles in China: how do we analyze them?* Washington DC: Presentation given at SAIS course. Miles, J. Special Report on China: Rising power, anxious State. *The Economist*, June 25, 2011, p. 11. The debt volumes of these companies are estimated at 14,000 billion RMB (about 2.1 trillion RMB). See Rabinovitch, S. Emphasis on explicit debt hides extent of Beijing's local liabilities. *Financial Times*, June 28, 2011.

4 Prospects of Property markets

Chart 5 illustrates the duration of a downturn of real housing prices from peak to bottoming out. On average, house prices after a downturn bottom out after 6 years. The decline in real house prices from peak to trough averages 35%.

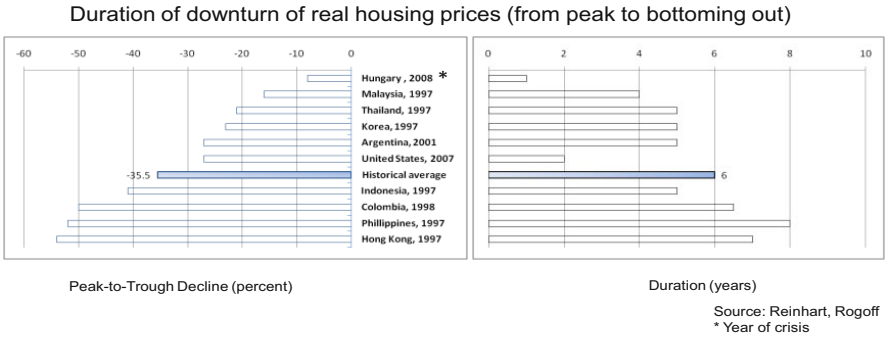


Chart 5: Duration of downturn of real housing prices

According to Chart 5, the most severe real housing declines were experienced in Colombia, the Philippines, and Hong Kong. Their crashes amounted to about 50% to 60% (from peak to trough). The housing price decline experienced by the United States in the aftermath of the financial and economic crisis of 2008/09 is at least two times stronger than the bust during the Great Depression in the 1930s.[24]

The long duration (on average 6 years) underlines the considerable impact of a housing crisis on economic development. After five years, an estimated increase of price rent ratios by 4% in 2011 and 2012 each indicate improved conditions in the property market as rising rents make home ownership more attractive.[25] A recovery of the housing market is looming.

The biggest stimulus for property markets is expected to come from emerging markets and is the result of rising investments in infrastructure (e.g. roads, transport systems, hospitals, etc.) and residential housing. Cities are expanding, not only in China but across most emerging markets, rural populations are migrating rapidly to expanding cities, increasing the demand for new homes, public infrastructure, factories, transportations, offices and shopping centers. By 2030, it is estimated that there will be more than 8 billion people on the earth, and three out of five will live in cities. At present, about 5 million people, through birth or migration, take up residency in urban areas every month.[26]

---

[24] Reinhart, C. M., & Rogoff, K. S. (2009). *This Time is Different – Eight Centuries of Financial Folly*. Princeton University Press: Princeton and Oxford, p. 226.

[25] The housing market: The darkest hour. *The Economist,* May 21, 2011 p. 32.

[26] Hurley, A. K. (2011). *(Re)conceiving cities.* One. John Hopkins Carey Business School, Volume III, No. 2, Spring/Summer 2011, 13-19, p. 13.

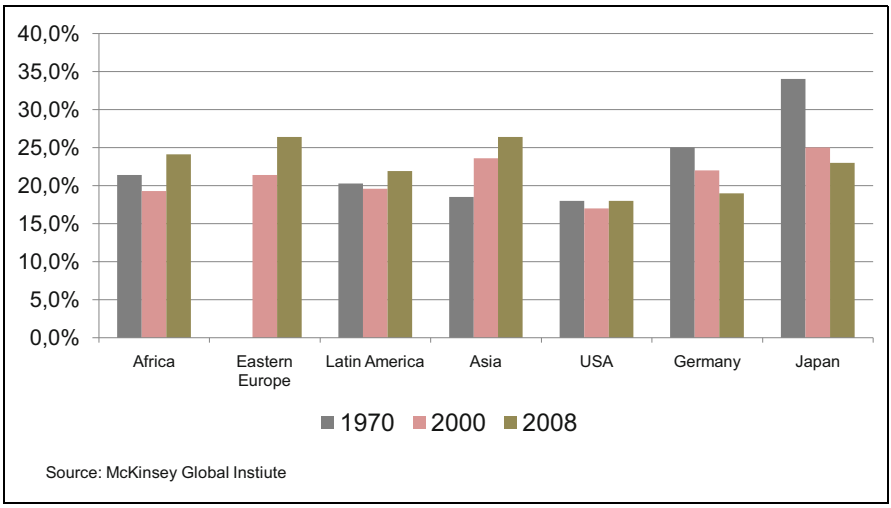


Chart 6: Investment to GDP in emerging markets versus in advanced economies

Additionally, the population in emerging market is relatively younger than in advanced economies, further propelling the demand for housing. For example, more than 60% of the roughly 300 million strong population of the Arab world is under the age of 25. One of their biggest concerns is the access to affordable housing.[27]

Chart 6 confirms the trend. Whereas investments in advanced economies have declined from 1970 to 2008, they have increased in emerging markets during the same period. For example, Africa has boosted its investment rate to 24.1% of GDP in 2008 from 19.3% in 2000. Eastern Europe's rate rose to 26.4% from 21.4% over the same period.[28]

Chart 7 provides an estimate of the potential volume to be invested into residential housing. According to estimates of the McKinsey Global Institute, global housing demand is projected to almost double over the next two decades from 2.1 trillion USD to about 4.0 trillion USD. Although this estimate also includes advanced economies, the major portion is absorbed by emerging markets.

[27] Drummond, J. Youths fear housing shortage. *Financial Times* April 1, 2011.

[28] Dobbs, R., Lund, S., et al. (2010). *Farewell to cheap capital? The implications of long-term shifts in global investments and savings*. McKinsey Global Institute: Seoul, San Francisco, London, Washington, p. 33.

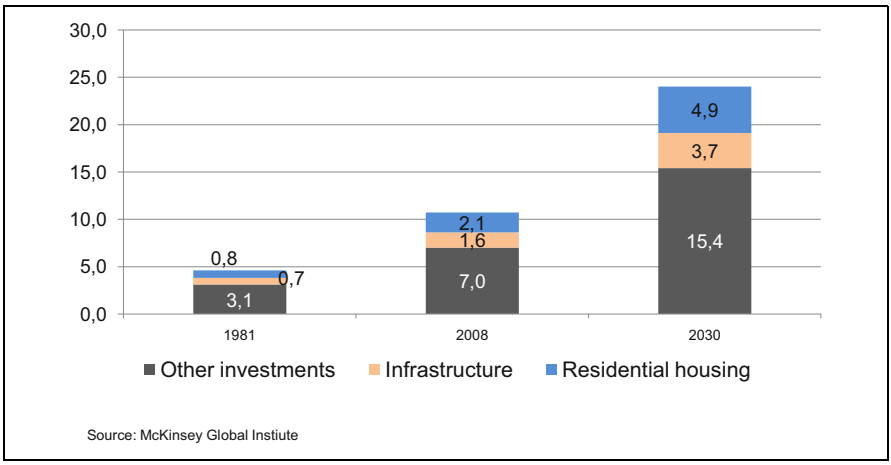


Chart 7: Estimated breakdown of investments globally in billion USD

Most developing countries currently invest less in residential real estate as a share of GDP than mature economies, reflecting their lower household income levels. India and Brazil for example, currently invest only 1.6% of GDP in residential estate (for comparison: for advanced economies, this number amounted on average to 4.6% in 2008). This ratio is likely to change, however, as incomes rise. In addition, increasing urbanization and a growing population require governments to act to avoid social unrest and political instability. The prospects for property markets are therefore benign, but major activities will shift from advanced economies to emerging markets.

Rising demand for investments is likely to push interest rates upwards and savings rates in emerging markets are expected to decline. Higher interest rates are likely to improve the economics of commercial and retail banking. Although credit volumes will likely grow more slowly as higher interest rates dampen loan demand and as a result of Basel III regulations,[29] net interest rate margins may increase as the cost of retail funding typically rises less than lending rates. Lenders may also rethink their business models in retail banking because in emerging markets households typically have lower incomes than in advanced economies.

5 Conclusion

Properties are the Achilles heel of the financial system. Housing bubbles often precede a banking crisis. Due to the structure of investments in property, they tend to reinforce cyclical effects. There is no difference between emerging markets and

[29] For detailed information about the changes in the Basel III framework in relation to the Basel II, see on the Bank for International Settlement's website (www.bis.org). A number of papers are ready for download.

advanced economies in the rise of property bubbles and the bust of the bubble. The duration of recovery of property markets is longer and more painful than a crash in the equity markets.

To ensure sound growth of housing finance markets, policy makers need to promote sounder primary mortgage markets: these should include LTV limits which are adjustable to the real estate market cycle, limitations on loan products which are capable of deteriorating affordability in case of macroeconomic shocks. For example, the granting of mortgage loans denominated in foreign currency should be linked to stress tests of the applicant's income to ensure that her/his income withstands an increase of the redemption rate due to fluctuations in exchange rates. Another important area is the development of house price indices and databases on property transactions to spot house price inflation. Coupled with these is the development of strong consumer protection frameworks which help consumers to understand better the risks they are exposed when they take out a mortgage loan.

A further challenge is the revival of capital market instruments to fund housing. The markets still suffer as investors have lost confidence in mortgage securities. Issuers face difficulties in issuing mortgage related debt.[30]

Due to rising demand for capital to finance the expected investment boom in emerging markets, the importance of globally integrated capital markets is also likely to rise, as domestic capital markets are unlikely to generate the necessary financing volumes. This demand provides opportunities for the development of new financing instruments, albeit more prudently structured due to their different capital requirement and a higher risk aversion of investors. On-balance funding instruments like covered mortgage bonds may become more attractive.[31] Thus, a number of opportunities will arise in emerging markets regarding the construction and financing of property.

Will house price bubbles still occur? Carmen Reinhart and Kenneth Rogoff point out that policy makers, bankers and borrowers are caught by the delusion of the "this time is different"-syndrome as early warning signs are overlooked when markets are growing. To draw an analogy, it is like "The Never-ending Story": "This is a different story which will be told another time".

[30] Hughes, J., & Van Duyn, A. Subprime hangs over securitized debt revival. *Financial Times*, June 10, 2011.

[31] A covered mortgage bond is a debt instrument which is secured against a dynamic pool of specifically identified, eligible mortgages. The fundamental concept of this security is the reliance on the collateral (mortgage) as the primary source of credit quality, which significantly reduces the risk to the bondholder. Mortgage bonds are issued by a bank and usually remain on its balance sheet. The credit quality of the bonds is assured through conservative underwriting standards and strict regulation of loans and lending institutions as well as strict valuation rules.

References

Clemens U., Dyck S., & Just, T. (2011). China's housing markets: Regulatory interventions mitigate risk of severe bust. Deutsche Bank Research: Frankfurt.

Dobbs, R., Lund, S. et al. (2010). Farewell to cheap capital? The implications of long-term shifts in global investments and savings. McKinsey Global Institute: Seoul, San Francisco, London, Washington.

Drummond, J. Youths fear housing shortage. Financial Times, April 1, 2011.

EMF Hypostat 2005 (2006), A review of Europe's Mortgage and Housing Markets, European Mortgage Federation: Brussels.

Ende, M., & Manheim, R. (1979). The Never-ending Story. K. Thienemanns Verlag: Stuttgart.

Hegedus, J ., & Somogyi, E. (2005). An Evaluation of the Hungarian Mortgage Program 2000–2004. In Struyk, R., & Hegedus, J. (Ed.), Housing Finance – New and Old Models in Central Europe, Russia and Kazakhstan. Open Society Institute: Budapest.

Hughes, J., & Van Duyn, A. Subprime hangs over securitized debt revival. Financial Times, June 10, 2011.

Hurley, A. K. (2011). (Re)conceiving cities. One. John Hopkins Carey Business School, Volume III, No. 2, Spring/Summer 2011, 13-19.

Miles, J. Special Report on China: Rising power, anxious State. The Economist, June 25, 2011.

Plender, J. Overarching Problems. Financial Times, January 27, 2011.

Rabinovitch, S. Emphasis on explicit debt hides extent of Beijing's local liabilities. Financial Times, June 28, 2011.

Rajan, R. G. (2010). Fault Lines – How hidden fractures still threaten the world economy. Princeton University Press: Princeton and Oxford.

Reinhart, C. M. & Rogoff, K. S (2009). This Time is Different – Eight Centuries of Financial Folly. Princeton University Press: Princeton and Oxford.

Renaud, B. (2011). Housing Bubbles in China: How do we analyze them? Washington DC: Presentation given at SAIS course.

Renaud, B. (2009). A Systemic View of Housing in China's New Urban Area. Seoul: Presentation at KDI School.

Rosenberg, M. (2010). China's population. http://geography.about.com/od/populationgeography/a/chinapopulation.htm.

Roxburgh, C., Lund, S. et al. (2009). Global capital markets: Entering a new era. McKinsey Global Institute: Washington, Seoul, San Francisco, London.

The Economist (2011), Bricks and Slaughter – A special report on property.

The Economist, May 27, 2011. Home Truths: China's property market.

The Economist, March 5, 2011. The housing market: The darkest hour.

Part B

Balanced Growth – a business perspective

Selecting the right growth mechanism: The choice between internal development, strategic alliances, and mergers & acquisitions

Markus Kreutzer

Abstract:

When seeking to realize growth strategies, firms have three choices: internal development, alliances, and mergers & acquisitions. How to choose between these growth mechanisms is, however, not well-understood in practice. Managers seldom sufficiently strategically analyze this critical first step in any growth-related decision process. Instead, many managers commonly base their decision on "gut-feel" or simply follow successful traits. This chapter offers a framework to systematically guide managers in their choice of growth mechanisms. Four sets of factors should be simultaneously considered to decide on when to make, ally, or buy: the environment, the target, the growth strategy, and company-related factors. By systematically analyzing each growth mechanism's context-specific advantages and disadvantages, firms may avoid mistakes that could not be compensated for in later implementation phases.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_6,

Table of Contents

1 Introduction

Most large organizations are well aware of the necessity to generate growth (e.g., Lechner and Kreutzer, 2010). If their growth stalls, they risk a string of negative effects, such as failure to implement economies of scale, to attract and retain key employees, or to raise sufficient capital (Goold, 1999).

However, realizing sustained, profitable corporate growth remains a challenging task for most firms (Lechner and Kreutzer, 2010). The sad truth is that many firms do not achieve their growth targets. Unsatisfactory results are generally ascribed to mistakes in three aspects: (1) content, (2) process, and (3) pace of growth. (1) *Growth content* refers to the growth agenda or strategy. At a very abstract level, firms can grow within their core business, adjacent to it by building on existing capabilities, or in totally new domains and businesses that are distant from their original core. Unsurprisingly, empirical evidence shows that firms fail more often the further they stray from their original core (Zook and Allen, 2003; Lechner, 2006). Lack of experience, skills, and capabilities in the new area are key reasons for this failure. (2) The *growth process* refers to the coordination of the implementation process within an organization. Lechner and Kreutzer (Lechner, Kreutzer et al., 2008; Lechner and Kreutzer, 2010; Kreutzer and Lechner, 2011), for example, differentiate between four modes of coordinating growth initiatives, which are distinguished by the degree of centralization, formalization, and informal social relationships. Furthermore, the authors show that a fit between the growth mode and the initiative content and context is crucial for success. (3) The *growth pace* refers to the optimal company growth rate. Empirical evidence shows that only those firms that grow within a so-called growth corridor, in which they exceed a certain required minimum growth but also avoid excessive growth, survive in the long-term (Raisch and von Krogh, 2007).

A fourth aspect is decisive for more or less successful growth ambitions, although managers, who tend to blame the implementation for goals not achieved, often overlook it (Capron and Mitchell, 2010): Selecting the optimal *mechanism* to realize growth strategies.

Managers have multiple growth mechanisms at their disposal. On a first level, two broad mechanisms can be differentiated: internal and external growth. Internal or organic growth builds on a firm's own development resources. However, in today's environments, it is increasingly unrealistic to only develop all required resources internally (Hoffmann and Schaper-Rinkel, 2001). External growth can "help bridge the gap between the firm's present resource endowment and future expected requirements" (Hoffmann and Schaper-Rinkel, 2001, p. 132), either by cooperation (strategic alliances), in which separate companies maintain separate identities, or by mergers and acquisitions (M&A), in which one firm acquires a controlling interest in another.

In short, firms have three choices when undertaking a growth initiative: make, ally, or buy (White, 2000; Geyskens, Steenkamp et al., 2006). Firms seem to be aware that these mechanisms represent alternative means: In a survey of US company executives, for example, 82% answered the question whether they view acqui-

sitions and alliances as two different ways of achieving the same growth goals in the affirmative (Dyer, Kale et al., 2004). However, few executives really treat acquisitions and alliances as alternative growth mechanisms in practice. This is evident from the same executives' responses to two further questions: Only 24% considered forming an alliance as an alternative the last time they executed an acquisition, and only 14% of the firms had specific guidelines and criteria in place to choose between growth mechanisms. Instead, firms mainly seem to follow experience; that is, due to prior positive acquisition experience, for example, they simply consider acquisition in the next growth initiative (Ernst and Halevy, 2000).

This is alarming if one considers that knowing when to use which mechanism can deliver – or fail to deliver – a competitive advantage (Dyer, Kale et al., 2004). A wrong or sub-optimal initial growth mechanism choice cannot be compensated in later implementation phases (Capron and Mitchell, 2010). In addition, empirical evidence suggests that firms that use all three alternative mechanisms outperform others (Capron and Mitchell, 2010).

It is crucial to ask which mechanism is most suited to a specific growth initiative. Research can offer guidelines and specific advice on when to opt for which mechanism (e.g., Hennart and Reddy, 1997; Ernst and Halevy, 2000; White, 2000; Hoffmann and Schaper-Rinkel, 2001; Roberts and Liu, 2001; Hagedoorn and Duysters, 2002; Dyer, Kale et al., 2004; Villalonga and McGahan, 2005; Geyskens, Steenkamp et al., 2006; Wang and Zajac, 2007; Yin and Shanley, 2008; Capron and Mitchell, 2009; Wiklund and Shepherd, 2009; Capron and Mitchell, 2010; Carayannopoulos and Auster, 2010). This literature stream on "governance choices" builds on several theoretical perspectives to answer this question. Originally, especially transaction cost economics (Williamson, 1985) and resource dependence theory (Pfeffer and Salancik, 1978) were used. More recently, other theories have complemented such analysis; these include the knowledge-based view (e.g., Carayannopoulos and Auster, 2010) and option theory (e.g., Folta and Miller, 2002). They all predict that organizations following the theoretical prescriptions and aligning organizational growth with the derived dimensions will achieve higher performance.

The dimensions are classified into four main sections: (1) environment, (2) target, (3) growth strategy, and (4) firm-related criteria. After highlighting the differences between the three growth mechanisms in Section 2, the dimensions are discussed in Sections 3-6, and the conclusion provides guidelines and recommendations in Section 7.

2 Growth mechanisms

2.1 Internal development

Internal development (or organic growth) is the first, and often still predominant, growth mechanism (Müller-Stewens, Menz et al., 2010). Firms can harness their competencies and capabilities to build a new business from scratch without relying

on external partners (Kling, Ghobadian et al., 2009). Innovation is a cornerstone of firms making extensive use of organic growth, such as 3M. This innovation is evident in high internal R&D expenditure and the desire to hire and retain the best employees (Rothaermel and Hess, 2010). Compared to external growth mechanisms, internal development is less risky but in most cases requires more time to develop the new business (Rothaermel and Hess, 2010).

2.2 Strategic alliances

Today, firms that only rely on internal growth are the exception. To swiftly and efficiently access external assets, resources, and knowledge in order to learn from others, two non-organic external growth mechanisms complement internal growth: strategic alliances and M&A. Strategic alliances are "voluntary agreements between [two or more] firms involving exchange, sharing, or co-development of products, technologies, or services" (Gulati, 1998, p. 293). They clearly differ from M&A in that the linked firms maintain legal and economic independence, as well as separate identities. In these agreements, joint decision-making is required (Bamford, 2003).

Strategic alliances can take on different forms, such as contractual alliances, equity joint ventures, or minority (non-controlling) equity stakes. Compared to internal development, they involve a higher risk of failure – about 50%, according to most studies (e.g., Bleeke and Ernst, 1995) – owing to potential competition between partners, negative spirals of mistrust, and a lack of management attention.

2.3 Mergers & acquisitions

In contrast to strategic alliances, firms lose legal and economic independence when they opt for mergers or acquisitions that secure a controlling majority (e.g., Hoffmann and Schaper-Rinkel, 2001). While M&A allow full control by establishing a clear hierarchy, signaling a high commitment, and usually ensuring managerial attention, investment costs – physical, human, and financial – can be significant. The reversibility of the growth strategy and strategic flexibility may therefore be decreased (e.g., Müller-Stewens and Brauer, 2009; Rothaermel and Hess, 2010). M&A activity failure rates are reported to be even higher than for strategic alliances. Over time, 65% of acquisitions have destroyed more value than they create (e.g., Laurie, Doz et al., 2006).

The advantages and disadvantages of each of the three mechanisms are summarized in Table 1.

Table 1: Comparison of the three growth mechanisms

	Internal development	Strategic alliances	Mergers & acquisitions
Characteristics and types	– Organic growth based on internal resources, skills, and capabilities – Innovation, R&D, and hiring and retaining the best human capital	– Firms maintain legal and economic independence/autonomy – Contractual alliances, joint ventures, or minority equity investments between at least 2 firms	– Firms lose legal and economic independence/ autonomy – Mergers or acquisitions (with either a majority or minority interest with controlling agreement)
Advantages	– low risk – reversibility – internalization of skills and capabilities – full capture of returns – better control of intellectual property	– shared risk – flexibility (e.g., small-scale investments provide strategic options; combination of multiple partners) – easier to exit if necessary; less separation pain – relatively swift	– full control; easier to establish hierarchical control – commitment and management attention – relatively swift
Disadvantages	– high investment (physical, human, and tangible resources) requirements – relatively slower – full risk exposure	– less capability to intensively exploit an opportunity; shared returns – threat of opportunistic behavior by partner(s) – challenge to align goals – potential competition between partners – negative spirals of mistrust	– high investment (physical, human, and tangible resources) requirements – high management costs – reduced flexibility – lower reversibility; higher separation pain and potential lock-in – risk of overpaying – cultural integration

3 Environmental criteria for mechanism selection

The first set of criteria that managers should consider for their choice between growth mechanisms is the environment, in terms of (1) external regulatory and institutional constraints and (2) environmental uncertainty.

3.1 Regulatory and institutional constraints

Firms' decisions regarding growth mechanisms have to conform to the common legal environment's prescriptions. Their choice may be restricted by institutional forces and political influences (Scott, 2001). Take foreign market entry as an example. In some countries, foreign acquisitions and greenfield investments are banned in selected (e.g., defense) or all sectors, or are made difficult by legal restrictions regarding voting rights or cross-holdings (Hennart and Reddy, 1997). Foreign automakers in China, for example, are only allowed to set up joint ventures with local partners, as China's government wants local firms to learn (Peng, 2009, p. 162). Another example is antitrust scrutiny in an industry. Yin and Shanley (2008) expect more alliances than M&A if an industry's recent antitrust activity has been high.

3.2 Environmental uncertainty

Environmental uncertainty, i.e. the turnover, absence of patterns, and unpredictability of the environment and of future payoffs (Dess and Beard, 1984), should influence the governance choice. In uncertain, fast-paced environments – in terms of both the technology/product/service and its acceptance by consumers –, alliances are more favorable than internal development or acquisitions. Internal development may take too long to impact the market and may need too many resources (Liebeskind, Oliver et al., 1996; Carayannopoulos and Auster, 2010). The time disadvantage is especially relevant in industries with steep learning curves and in network industries in which customer lock-in is key (Müller-Stewens and Brauer, 2009).

Acquisitions may be too risky in uncertain environments because of their high resource investments and high exposure. Unforeseen changes may negate prior investments' potential value add (Yin and Shanley, 2008). Besides lower initial investment needs, strategic alliances provide two additional advantages in uncertain environments: the potential for (enhanced and faster) learning and flexibility, as they are easier to set up, reverse, and exit (Hoffmann and Schaper-Rinkel, 2001; Yin and Shanley, 2008). By collaborating, firms can limit their exposure, while simultaneously maintaining the option to invest more in the alliance if it begins to show results or to acquire the partner (Dyer, Kale et al., 2004).

Many firms deliberately interpret their alliances as strategic options. ABInBev, for example, a global player in the beverage industry, and Cisco, a global networking firm, like to work with a company by means of an alliance for at least one year before they decide whether an acquisition will work. Similarly, almost half of the global pharmaceutical firm Eli Lilly's acquisitions involve prior partners (Capron and Mitchell, 2010). Bleeke and Ernst (1993) show that 75% of all alliances end with one of the partners acquiring the other. Multipartner alliances offer even more flexibility and speed advantage and are especially attractive for setting standards (e.g., Ernst and Halevy, 2000; Lavie, Lechner et al., 2007). Acquisitions of more than two firms would be too costly and disruptive.

In line with the abovementioned prediction, prior empirical studies find that governance choices vary between dynamic and more mature industries. Technological volatility, for example, increases alliance formation (e.g., Hagedoorn, 1993; Eisenhardt and Schoonhoven, 1996; Hagedoorn and Duysters, 2002). Recent evidence from Switzerland also shows that the application frequencies of the two external growth mechanisms differ between industries. Müller-Stewens, Menz, and colleagues (2010) found that manufacturing firms seek to increase their use of M&A, while life science firms seek to increase their use of alliances. In their study on alliance announcements' effects on share prices, Ernst and Halevy (2000) found that alliances are better received than M&A in fast-moving, highly uncertain industries such as electronics, mass media, and software.

4 Target-related criteria for mechanism selection

The second set of criteria refers to target-related aspects and, thus, applies more to the choice between the two external growth mechanisms.

4.1 Existence of potential partners

An obvious first question is whether there are any potential target firms to cooperate with or acquire that possess the necessary assets, skills, or capabilities (e.g., Hennart and Reddy, 1997). Daimler, for example, relied mainly on alliances in its recent internationalization process in the BRIC countries, forming a joint venture with Foton in China and a strategic cooperation with KAMAZ in Russia. In India, however, Daimler had to abandon this process as it could not find a suitable local partner.

4.2 Competition level

In addition, to decide on whether to ally with or acquire a potential partner, firms should validate the competition level around the target. Dyer and colleagues (2004) conclude that if any competitors are interested in the same target, a company may have to buy it right away to pre-empt the competition – at least, if the environmental uncertainty is not high.

Growth strategy timing is closely related to the first two target-related aspects. In another automotive example, it may be too late for an automobile firm to now expand into China by means of a joint venture (therefore, at all), since all (strategically, financially, and culturally) suitable and capable local partners are already in agreements with competitors.

4.3 Targeted assets' digestibility

Furthermore, firms should understand the degree to which a potential partner's assets are linked to or embedded in its non-desired assets. A firm will favor acquisitions over alliances when the assets it needs are not mixed with other non-desired assets and can thus be acquired by buying the firm or part of it (Hennart, 1988). Conversely, if assets are packaged in a complex way, such as in larger firms with a non-divisional structure, Hennart and Reddy (1997) find that joint ventures will be the chosen option.

4.4 Ease of assessing the target firm's value

Can you easily assess the assets in the target firm that interests you? This is another key question to pose (Capron and Mitchell, 2010). If this is not possible, or if you don't understand the target's business well enough, strategic alliances (joint ventures) may be most efficient, as you may gain more information over time and retain high flexibility in the interim (Balakrishnan and Koza, 1993).

4.5 Risk of opportunistic behavior

Finally, what is the likelihood that your potential partner may act opportunistically? The greater your perceived fear that your partner's behavior will be opportunistic, the stronger your need for control, and the more strongly you should consider acquisition rather than an alliance (Hoffmann and Schaper-Rinkel, 2001). Internal development may be the other alternative to gain full control.

5 Growth strategy criteria for mechanism selection

Besides considering the potential target's environment and characteristics, firms should be guided by the growth strategy in their choice of governance. Specifically, the growth strategy's (1) distance to the core, (2) the cultural and geographical distance, and (3) the potential for synergies should be included in the analysis.

5.1 Distance to the core

If firms grow incrementally in their existing core, it may be possible to leverage existing skills and capabilities. In this case, internal development is the most effective and fastest mechanism. By applying and slightly adapting internal skills to the new business, firms can enhance their capabilities and may benefit from a more motivated workforce, as they signal entrepreneurial spirit and innovation potential (Müller-Stewens and Brauer, 2009, p. 314).

If a firm's existing resources and capabilities are less relevant for the intended growth initiative, it should opt for external growth (Capron and Mitchell, 2010). If these needed capabilities relate to the firm's core business, it should opt for M&A, which guarantees the extent of the control it needs to save its core, as well as the high value involved. Acquisitions also offer swift entry into the highly valued knowledge sector (Carayannopoulos and Auster, 2010). If the needed capabilities relate to non-core businesses, alliances are ideal (Hagedoorn and Duysters, 2002).

The network firm Cisco applies this logic. When deciding on the most appropriate growth mechanisms, Cisco's corporate development division first establishes whether or not the technology and expertise exist in-house. If so, internal growth is the preferred option. If not, and the target is (not) intended to be part of the future core, it opts for an acquisition (alliance). Similarly, many firms – for instance, Arla Foods – apply the rule of not allying in respect of their core business.

5.2 Cultural and geographical distance

Existing classifications of growth strategies often subsume entering new markets (i.e. internationalization) as growth within or adjacent to the core (Zook and Allen 2003; Lechner, 2006). However, such growth initiatives may involve distance in terms of culture and geography, even though the same product range is involved. On average, a German firm entering a new market faces a higher distance when entering the Indian market than it does when entering the Austrian market. The

higher the distance in both dimensions, the better alliances work when compared to the alternative mechanisms.

In line with this argument, Ernst and Halevy (2000), who studied alliance announcements' effects on share prices, find alliances to be the preferred choice for companies seeking to enter new geographies. On the other hand, Kogut and Singh (1988) find that joint ventures are preferred to acquisitions, the greater the cultural distance between two firms is. They cite the potential management costs involved in integrating the target firm's workforce after an acquisition as one of the reasons for the preference. In practice, leading firms recognize this criterion. Cisco, for example, seeks to avoid acquisitions if employees would have to move far, or if vital facilities are not located nearby.

Again, having the potential to build on existing skills and capabilities changes the picture. When you have both solid market knowledge and the right distribution channels (Gulati, 2004) or capabilities to manage a foreign labor force (Hennart and Reddy, 1997), acquisitions can be a good alternative when entering a new geographic location. Siemens, for example, expanded swiftly into major electronics markets such as the US by means of M&A (Gulati, 2004).

5.3 Synergy potential

According to Dyer and colleagues (2004), the interdependencies between potential partners should also guide the governance choice. The more the growth strategy requires close cooperation and iterative knowledge-sharing processes (i.e. reciprocal synergies), the higher the required degree of integration and the better acquisitions work, as they already offer early-stage synergy realization potential. However, it must be noted that such synergy potential is often overestimated.

The scope of the growth initiative also influences synergy realization. If executing the growth strategy requires a broad scope of joint activities and involves many people and units, M&A may be a better solution than a partnership to realize synergies at full control (Capron and Mitchell, 2010).

In addition, the nature of the resources to be combined is relevant. Are they "harder" (e.g., manufacturing plants) or "softer" (e.g., knowledge and human capital) (Dyer, Kale et al., 2004)? Dyer and colleagues argue that the higher hard assets' relative importance, the better acquisitions work. When mainly soft assets have to be combined, M&A should be avoided. Evidence from the pharmaceutical industry corroborates this logic. Large pharmaceutical firms (such as Roche) predominantly opt for licensing agreements and other types of strategic alliances to gain access to promising drugs from small, innovative biotech firms to fill their drug pipelines. If they acquired the small biotech firms, they would risk losing key employees (scientists) in the integrated firms who dislike working for large organizations with strong hierarchies.

6 Company resources and capabilities criteria for mechanism selection

As a fourth and final set of criteria, firms' own resources, skills, and capabilities are relevant. Specifically, when deciding how to grow, managers should consider their firm's (1) existing resource basis, (2) growth mechanism capabilities, and (3) skills in absorbing and protecting knowledge. Company-specific factors may account for the observation that there are not only interesting differences between sectors but also between competitors in the same industry (e.g., in the computer software industry, between SAP and Oracle) (Müller-Stewens and Brauer 2009, p. 316).

6.1 Resource endowments

Managers should consider whether they are endowed with relevant resources and have the financial strength to pursue resource-intensive growth mechanisms such as internal development and M&A. Unsurprisingly, Hoffmann and Schaper-Rinkel (2001) find support for the hypothesis that financially stronger firms favor acquisitions over alliances.

6.2 Growth mechanism capabilities

Evidence suggests that firms follow successful traits. Several studies (e.g., Villalonga and McGahan, 2005; Carayannopoulos and Auster, 2010) confirm that a firm's prior experience with alliances or acquisitions predicts the external growth mechanism's future choice. Unsurprisingly, recent success with one mechanism further increases the likelihood of its subsequent adoption (Haleblian, Kim et al., 2006). For example, if a firm acquired successfully in the past, it will be more likely to opt for an acquisition to realize the next growth initiative (Ernst and Halevy, 2000). This may be a good thing, because experience pays off. We know that firms can build distinct capabilities in one growth mechanism, for example, an alliance capability (e.g., Kale, Dyer et al., 2002; Schreiner, Kale et al., 2009) and are – on average – more successful at applying this mechanism in future. However, "a company's initial experience often turns into blinders" (Dyer, Kale et al., 2004, p. 110) and endangers its long-term success. Insisting on using one mechanism only, even if a company is very good at it, only adds value if all other circumstances demand this mechanism's use.

Evidence suggests that, after making their first acquisition, especially relatively inexperienced acquirers inappropriately generalize acquisition experience to subsequent dissimilar acquisitions (i.e. dissimilar growth strategies) (Haleblian and Finkelstein, 1999). The converse may be equally risky: Firms that exclude one mechanism from consideration owing to disappointing past experiences. For instance, a German automotive supplier opted for a joint venture in the Indian market. While this choice was fully in line with the argumentation presented here, subsequent decisions regarding the partner selection and alliance governance structure caused the initiative to fail. As a result of this failure, the firm did not consider the option to enter the Chinese market by forming an alliance.

6.3 Absorptive capacity and appropriability regime

Managers should also assess their firm's ability to protect their unique skills and innovative technologies. The stronger their firms' appropriability regime (Hagedoorn and Duysters, 2002), the more shielded it is from quick imitation and opportunistic partner behavior. Thus, alliances become a less risky alternative.

The more and the faster a firm is able to learn and adapt in an alliance, i.e. the higher its absorptive capacity (Cohen and Levinthal, 1990), the higher the chances to "outlearn" partners in so-called "learning races" (Hamel 1991; Khanna, Gulati et al., 1998). Generally, the higher the absorptive capacity, the lower the need to control and, thus, the more benefits can be derived from an alliance (Hoffmann and Schaper-Rinkel, 2001).

7 Discussion and recommendations

This chapter started off with the observation that managers seldom sufficiently strategically analyze the critical first step in any growth-related decision process: the governance choice. Instead, most managers base their decisions to realize growth strategies by internal development, strategic alliances, or M&A on "gut feel." A framework was presented that may systematically guide managers in their choice by considering four sets of factors: environmental, target, growth strategy, and company-related factors. A firm's *make, ally, or buy* decision should be an outcome of managers' simultaneous consideration of all the dimensions applicable to a specific context (for an overview, see Table 2).

In short, for an acquisition to be the most suitable option, there must be, for example, no legal constraints, a low to medium degree of market uncertainty should prevail, and first movers should be rewarded. There must be potential targets that are not highly competed on, and the targeted assets should be easy to digest and assess. Intended growth should focus on the current core business or not-too-distant future core business, and should have the potential to realize reciprocal synergies, mainly through hard assets. And the company itself should ideally be able to build on in-house M&A capabilities and a solid financial basis.

Table 2: Guidelines on when to use which growth mechanisms

Criteria related to	The following mechanisms are: ++ = highly effective, + = effective, -- = not effective at all, when	Internal	Strategic alliances	M&A
Environment	All growth mechanisms are legally possible and not constrained by any institutional norms.	++	++	++
	Market uncertainty in terms of products and technology is high.	+	++	--
	Market uncertainty in terms of consumer acceptance of new products/services is high.	+	++	--
	First mover advantages arise in the targeted business segment.	--	+	++
Target	There are no potential partners.	++	--	--
	The competition level for potential targets is high.	+	--	++
	The targeted assets are hard to digest.	+	++	--
	The target firm's value cannot be easily assessed.	+	++	--
	The risk of opportunistic behavior is high.	+	--	++
Growth Strategy	The company grows in its core and can build on its existing skills and capabilities.	++	--	+
	The company grows in its core but is dependent on external skills and capabilities.	--	--	++
	The company grows outside its core and is dependent on external skills and capabilities.	--	++	+
	Cultural distance is high.	--	++	--
	Geographical distance is high.	--	++	--
	The potential for reciprocal synergies is high.	++	--	++
	The intended scope of the growth strategy is high.	++	--	++
	The relative value of soft to hard assets is high.	+	++	--
Firm	The company's financial strength is high.	++	--	++
	The company has an innovation capability.	++	n.a.	n.a.
	The company has a strategic alliance capability.	n.a.	++	n.a.
	The company has an acquisition capability.	n.a.	n.a.	++
	The company's ability to protect its critical knowledge is high (high appropriability regime).	+	++	+
	The company's ability to learn and absorb new knowledge is high (high absorptive capacity).	+	++	+

Having outlined this framework, it is important to note several aspects. First, as one might imagine, in some cases, no clear-cut solution results from applying this framework. However, such a decision process may help structure organizational decision-making, reveal underlying assumptions, and point to potential conflict or different interpretations. By weighting the dimensions, firms may adapt the framework to their context. For example, the assessment of environmental uncertainty may overrule the level of competition for a certain target firm in the sense that, even

if there is high competition for a target, one should refrain from buying it right away if environmental uncertainty is high. Firms should conduct a cost-benefit analysis and opt for the governance choice with the highest net gain (Yin and Shanley, 2008). Structurally, responsibilities for all three growth mechanisms should be combined in one position to realize this – as Cisco does, for example, with its senior vice president for corporate development (Dyer, Kale et al., 2004).

Second, this governance choice is only a first, high-level step. In practice, the available governance options are more diverse and fine-grained. Deciding to ally is not enough; firms must choose from a variety of collaboration structures (Ernst and Halevy, 2000), such as an equity joint venture or a contractual alliance. Again, such choices are case-specific and context-specific. The underlying rationales presented here, especially the trade-off between flexibility and control, also apply to this more fine-grained choice. In the market, there is a general trend towards simpler and more flexible deal structures (Ernst and Halevy, 2000).

Third, the clear boundaries drawn between the three mechanisms are blurring. Many emerging multinationals (e.g., Tata) but also Western companies (e.g., Disney or Renault) recently applied a partnering approach to M&A; that is, they deliberately keep an acquisition separate and maintain its identity and organizational structure (Kale and Singh, 2009).

Fourth, the choice between the three mechanisms should not be a once-and-for-all decision (Capron and Mitchell, 2010). Managers should regularly revisit the chosen growth mechanism by assessing the criteria. For example, the environment may have become more stable and may therefore allow acquiring a partner (thus, exercising the option).

Fifth, managers should also be aware that the governance choice, while a decisive decision, is only the first step in a subsequent decision and implementation process. If a firm decides that pursuing a strategic alliance is, for example, the most suitable option, the subsequent phases, which include a partner selection (e.g., Holmberg and Cummings, 2009), negotiation (e.g., Reuer and Arino, 2002), governance (e.g., Reuer and Arino, 2007), management (e.g., De Man and Roijakkers, 2009), assessment, and termination (e.g., Gulati, Sytch et al., 2008) are key to achieving set growth targets. The initial mechanism choice only lays the foundation for successful company growth. Only if firms excel in execution, can the right governance choice pay off in the long term. And companies must invest in developing the ability to execute all mechanisms well if they want to grow profitably in the long term.

Sixth, firms that stick to the one growth mechanism they know may, as noted, be successful in the short term, but they probably "end up doing the wrong things really well, lagging more broadly capable competitors, and, quite likely, becoming targets themselves" (Capron and Mitchell 2010, p. 107). Thus, it is reasonable to assume that most mid-sized to large firms will ideally pursue their growth strategies with different growth mechanisms – as evidenced, for example, by Tesco's successful growth history (Kreutzer and Lechner, 2011).

Seventh, recent research has identified whether there are any positive or negative interactions between growth mechanisms in a single firm's portfolio. Rothaer-

mel and Hess (2010) analyzed the effects of simultaneously pursuing various mechanisms in global pharmaceutical companies that built innovative capabilities in biotechnology. Among others, they find synergies between internal R&D expenditures and alliances, suggesting that a strong internal R&D capability allows a company to select and pursue the most promising strategic alliances. Furthermore, alliances and acquisitions frequently reinforce one another. Overall, managers must understand and predict how these mechanisms are likely to interact in their organization when used in tandem.

Finally, future research – both theory development and empirical studies – should combine theoretical perspectives to explain the growth mechanism choice. Any single theoretical perspective – for instance, transaction cost economics – is empirically inconclusive (David and Han, 2004), ignores relevant factors such as company governance mechanisms' capabilities (Carayannopoulos and Auster, 2010), and does not account for multiple organizational levels.

References

Balakrishnan, S., Koza, M. P. (1993). Information asymmetry, adverse selection and joint-ventures – theory and evidence. *Journal of Economic Behavior & Organization* 20(1), 99-117.

Bamford, J. D. (2003). *Master alliance strategy: A comprehensive guide to design, management, and organization*. San Francisco, CA: Jossey-Bass.

Bleeke, J., Ernst, D. (1993). *Collaborating to compete*. New York, NY: John Wiley & Sons.

Bleeke, J., Ernst, D. (1995). Is your strategic alliance really a sale. *Harvard Business Review* 73(1), 97-105.

Capron, L., Mitchell, W. (2009). Selection capability: How capability gaps and internal social frictions affect internal and external strategic renewal. *Organization Science* 20(2), 294-312.

Capron, L., Mitchell, W. (2010). Finding the right path. *Harvard Business Review* 88(7-8), 102-107.

Carayannopoulos, S., Auster, E. R. (2010). External knowledge sourcing in biotechnology through acquisition versus alliance: A kbv approach. *Research Policy* 39(2), 254-267.

Cohen, W. M., Levinthal, D. A. (1990). Absorptive-capacity – a new perspective on learning and innovation. *Administrative Science Quarterly* 35(1), 128-152.

David, R. J., Han, S. K. (2004). A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal* 25(1), 39-58.

De Man, A.-P., Roijakkers, N. (2009). Alliance governance: Balancing control and trust in dealing with risk. *Long Range Planning* forthcoming.

Dess, G. G., Beard, D. W. (1984). Dimensions of organizational task environments. *Administrative Science Quarterly* 29(1), 52-73.

Dyer, J. H., Kale, P., Singh, H. (2004). When to ally and when to acquire. *Harvard Business Review* 82(7-8), 108-115.

Eisenhardt, K. M., Schoonhoven, C. B. (1996). Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. *Organization Science* 7(2), 136-150.

Ernst, D., Halevy, T. (2000). When to think alliance. *McKinsey Quarterly*(4), 46-55.

Folta, T. B., Miller, K. D. (2002). Real options in equity partnerships. *Strategic Management Journal* 23(1), 77-88.

Geyskens, I., Steenkamp, J., Kumar, N. (2006). Make, buy, or ally: A transaction cost theory meta-analysis. *Academy of Management Journal* 49(3), 519-543.

Goold, M. (1999). The growth imperative. *Long Range Planning* 32(1), 127-129.

Gulati, R. (1998). Alliances and networks. *Strategic Management Journal* 19(4), 293-317.

Gulati, R. (2004). How CEOs manage growth agendas. *Harvard Business Review* 82(7-8), 124-132.

Gulati, R., Sytch, M., Mehrotra, P. (2008). Breaking up is never easy: Planning for exit in a strategic alliance. *California Management Review* 50(4), 147-163.

Hagedoorn, J. (1993). Understanding the rationale of strategic technology partnering – inter-organizational modes of cooperation and sectoral differences. *Strategic Management Journal* 14(5), 371-385.

Hagedoorn, J., Duysters, G. (2002). External sources of innovative capabilities: The preference for strategic alliances or mergers and acquisitions. *Journal of Management Studies* 39(2), 167-188.

Haleblian, J., Finkelstein, S. (1999). The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Administrative Science Quarterly* 44(1), 29-56.

Haleblian, J. J., Kim, J. Y. J., Rajagopalan, N. (2006). The influence of acquisition experience and performance on acquisition behavior: Evidence from the US commercial banking industry. *Academy of Management Journal* 49(2), 357-370.

Hamel, G. (1991). Competition for competence and inter-partner learning within international strategic alliances. *Strategic Management Journal* 12, 83-103.

Hennart, J. F. (1988). A transaction costs theory of equity joint ventures. *Strategic Management Journal* 9(4), 361-374.

Hennart, J. F., Reddy, S. (1997). The choice between mergers/acquisitions and joint ventures: The case of Japanese investors in the united states. *Strategic Management Journal* 18(1), 1-12.

Hoffmann, W. H., Schaper-Rinkel, W. (2001). Acquire or ally? – a strategy framework for deciding between acquisition and cooperation. *Management International Review* 41(2), 131-159.

Holmberg, S. R., Cummings, J. L. (2009). Building successful strategic alliances strategic process and analytical tool for selecting partner industries and firms. *Long Range Planning* 42(2), 164-193.

Kale, P., Dyer, J. H., Singh, H. (2002). Alliance capability, stock market response, and long-term alliance success: The role of the alliance function. *Strategic Management Journal* 23(8), 747-767.

Kale, P., Singh, H. (2009). Managing strategic alliances: What do we know now, and where do we go from here? *Academy of Management Perspectives* 23(3), 45-62.

Khanna, T., Gulati, R., Nohria, N. (1998). The dynamics of learning alliances: Competition, cooperation, and relative scope. *Strategic Management Journal* 19(3), 193-210.

Kling, G., Ghobadian, A., O'Regan, N. (2009). Organic growth and shareholder value: A case study of the insurance industry. *International Journal of Research in Marketing* 26(4), 276-283.

Kreutzer, M., Lechner, C. (2011). Strategische Initiativen und Programme mit Führungsansätzen steuern. In M. Menz, T. Schmid, G. Müller-Stewens and C. Lechner *Strategische Initiativen und Programme. Wege erfolgreicher Unternehmensführung*. (69-88). Zürich: SGO.

Kreutzer, M., Lechner, C. (2011). Tesco versus Sainsbury's growth strategies and corporate competitiveness 1990-2007. In H. W. Volberda, R. E. Morgan, P. Reinmoeller et al. *Strategic management: Competitiveness and globalization (concepts and cases).* (795-813). EMEA: Cengage Learning.

Laurie, D. L., Doz, Y. L., Sheer, C. P. (2006). Creating new growth platforms. *Harvard Business Review*.

Lavie, D., Lechner, C., Singh, H. (2007). The performance implications of timing of entry and involvement in multipartner alliances. *Academy of Management Journal* 50, 578-604.

Lechner, C. (2006). Lernen, Wachstum zu managen. *io new management*(4), 24-28.

Lechner, C., Kreutzer, M. (2010). Coordinating growth initiatives in multi-unit firms. *Long Range Planning* 43(1), 6-32.

Lechner, C., Kreutzer, M., Schädler, J. (2008). Wachstumsinitiativen erfolgreich realisieren. *Harvard Business Manager* (Oktober), 74-84.

Liebeskind, J. P., Oliver, A. L., Zucker, L., Brewer, M. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. *Organization Science* 7(4), 428-443.

Müller-Stewens, G., Brauer, M. (2009). *Corporate strategy & governance. Wege zur nachhaltigen Wertsteigerung im diversifizierten Unternehmen.* Stuttgart: Schäffer-Poeschel.

Müller-Stewens, G., Menz, M., Lochbrunner, V., Gentle, C., Kobler, D., Grampp, M. (2010). The future for business in Switzerland? How to stay ahead. St.Gallen/ Zürich, University of St.Gallen/Deloitte.

Peng, M. W. (2009). *Global strategic management*. Mason, Ohio: South-Western/Cengage Learning.

Pfeffer, J., Salancik, G. (1978). *The external control of organizations.* New York, NY: Harper & Row.

Raisch, S., von Krogh, G. (2007). Navigating a path to smart growth. *MIT Sloan Management Review* 48(3), 65-72.

Reuer, J. J., Arino, A. (2002). Contractual renegotiations in strategic alliances. *Journal of Management* 28(1), 47-68.

Reuer, J. J., Arino, A. (2007). Strategic alliance contracts: Dimensions and determinants of contractual complexity. *Strategic Management Journal* 28(3), 313-330.

Roberts, E. B., Liu, W. Y. K. (2001). Ally or acquire? How technology leaders decide. *MIT Sloan Management Review* 43(1), 26-34.

Rothaermel, F. T., Hess, A. M. (2010). Innovation strategies combined. *MIT Sloan Management Review* 51(3), 13-15.

Schreiner, M., Kale, P., Corsten, D. (2009). What really is alliance management capability and how does it impact alliance outcomes and success? *Strategic Management Journal* 30(13), 1395-1419.

Scott, W. R. (2001). *Institutions and organizations*. Thousand Oaks, CA: Sage.

Villalonga, B., McGahan, A. M. (2005). The choice among acquisitions, alliances, and divestitures. *Strategic Management Journal* 26(13), 1183-1208.

Wang, L., Zajac, E. J. (2007). Alliance or acquisition? A dyadic perspective on interfirm resource combinations. *Strategic Management Journal* 28(13), 1291-1317.

White, S. (2000). Competition, capabilities, and the make, buy, or ally decisions of Chinese state-owned firms. *Academy of Management Journal* 43(3), 324-341.

Wiklund, J., Shepherd, D. A. (2009). The effectiveness of alliances and acquisitions: The role of resource combination activities. *Entrepreneurship Theory and Practice* 33(1), 193-212.

Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. New York, NY: Free Press.

Yin, X. L., Shanley, M. (2008). Industry determinants of the „Merger versus alliance“ Decision. *Academy of Management Review* 33(2), 473-491.

Zook, C., Allen, J. (2003). Growth outside the core. *Harvard Business Review* 81(12), 66-73.

Challenges and success factors for sustainable growth: Experience from strategy consulting

Jens Schädler and Melanie Oschlies

Abstract:

Economic recovery following a financial crisis includes companies in most sectors showing increased growth ambitions. Although some of the immediate growth opportunities might be tempting, a clear strategic path must be followed to achieve sustainable success. It is important to first define and assess a company's core business. The next steps are then – depending on the outcome of the first analyses – to focus on and realize the existing core business's full potential, to expand into adjacencies, or to redefine the core business. This chapter illustrates the way to become a sustainable value creator by raising a set of questions each manager must answer along a successful growth path. Potential answers to the individual questions are described, supported by selected company examples.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_7,

Table of Contents

1 The odds of sustainable growth strategies

Having emerged from the recent recession, European companies are actively pursuing growth strategies – organically, as well as inorganically. However, these growth attempts are challenging and often fail. Our studies have shown that only approximately 10% of companies are sustained value creators. Such companies create sustainable shareholder value through long-term returns above their cost of capital and achieve profitable growth on the top and bottom line.

In contrast, companies that fail in their growth attempts often lack a clear, ex ante core business definition, follow temptations blindly, or are too slow to implement their strategy by mobilizing the organization. Furthermore, companies frequently underestimate the importance of an industry leadership position in their core business. Managers also repeatedly fall into the trap of believing that they should focus on the most attractive markets, independent of their own market position. However, our research shows that 20-30% of companies with a leadership position in comparatively less attractive markets are sustained value creators, while only 2-8% of followers in highly attractive markets achieve this position. Clearly, leadership positions matter more than the market's attractiveness.

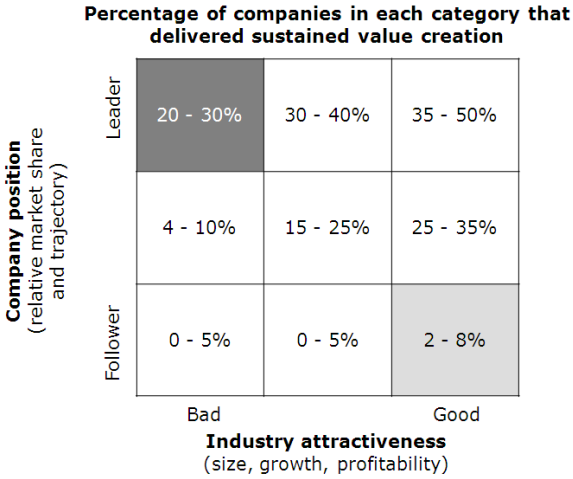


Figure 1: Sustained value creators (Source: Bain & Company)

To avoid a growth strategy failing, a clear path must be followed when developing such a strategy. Critical steps for building the basis of a successful growth plan are defining and assessing the core business. Once this assessment has been done, three different paths can be followed for sustainable growth, depending on the core business's state. Starting with the two initial steps, the process of formulating a sustainable growth strategy is described in this chapter. The success factors and challenges are highlighted regarding focus, expansion, and redefinition strategies that ultimately lead to a sustainable value creation.

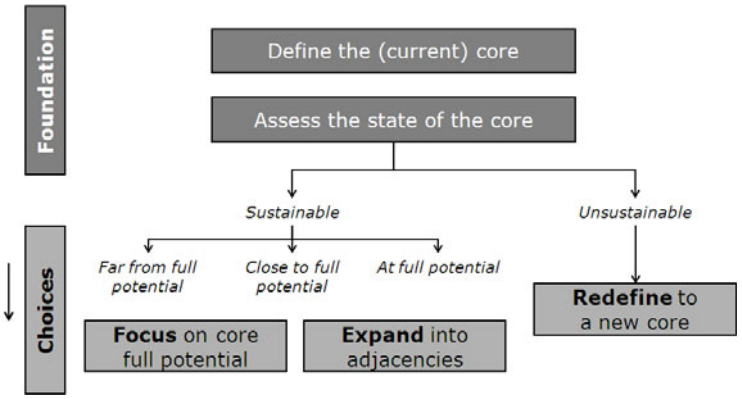


Figure 2: Process for growth strategy development (Source: Bain & Company)

2 The critical first steps

Critical first steps that are often neglected in growth strategies are the formulation of a crystal-clear core business definition and the subsequent analysis of the core business's current state. By following these two steps, companies can avoid growth attempts' most common pitfalls, such as spreading resources too thinly, confusing the organization or investors, and pursuing unsuitable adjacency paths. Therefore, before undertaking a growth initiative, managers need to ensure they can answer each of the following six questions.

2.1 Define the (current) core

What are our core markets?
How do we differentiate our business from those of our competitors?
What does our broader activity system look like?

To define a company's core, it is essential to review its core markets, its differentiating assets and capabilities, and the overall activity system. Furthermore, it is critical to consider the right economic boundaries and to have a clear business definition on which the entire management team has consensus. Management team members' lack of discussion and agreement on the core business is a frequent complication in the core business definition phase.

Polaroid and Gillette are examples of companies that had a too narrow or too broad core market definition with unclear boundaries. Polaroid defined its core market as "chemical/paper-based processing" and missed the broader definition of image capturing brought about by the digital age and ultimately went bankrupt. In contrast, Gillette defined its core business as "checkout purchases" and added batteries and writing instruments to its portfolio by acquiring companies in these segments. This definition was too broad and led to the organization's share price declining by more than 50% over three years.

Three criteria should be taken into account to successfully identify a company's core markets: customer loyalty, the potential for competitive outperformance, and market profitability. Customer loyalty and competitive outperformance are especially crucial, since more than 90% of sustained value creators are leaders in their core business and have twice as much customer loyalty as their competitors. A leadership position allows a better comparative cost position due to the accumulated experience and scale effects which, in return, lead to higher profits – even if the prices are similar to those of competitors.

Once the core markets have been clarified, the assets and/or capabilities that define the company's positioning in relation to its competitors and enable a leadership position need to be identified. These core assets and capabilities should be unique, inimitable, tangible, differentiable, and mutually reinforcing to gain long-term competitive advantage (see Barney, 1991). The final step is to anchor the core assets and capabilities within the company's business model.

2.2 Assess the state of the core

Is the core sustainable?
Is the core's state close to its full potential?
What are the strategic implications?

In order to become a sustainable value creator, it is important to assess whether the identified core business can generate profits for at least five more years without a major redefinition. This can be done by analyzing the robustness of the differentiation sources against competitors' models, technological trends, and the expected shifts in the industry's profit pool. Besides providing insights into the sustainability of a company's core activities, this assessment should also be used to look for unidentified growth opportunities. There is often substantial untapped potential in a company's core business that could allow for profitable growth in core markets with the current capabilities and assets and does not require adjacent growth activities. Looking for these growth opportunities are thus valuable time and resource investments. Potential areas of untapped growth potentials could be customers, the cost base, and the competitive positioning. Indicators of such areas are, for example, a decline in the core clients' share of wallet, a low or decreasing customer advocacy, increased business, product, and process complexity, and an increase in competitors' market shares and/or reinvestment rates.

When the foundation has been laid by answering the first six questions – from the core business definition to the status assessment – it is time to choose the future path. As illustrated in Figure 2, if the core business is sustainable but below its full potential, it is time to focus (see Section 3). If the core business is sustainable and at full potential, it is time to expand (see Section 4). Finally, if the core is unsustainable, a redefinition might be the right choice (see Section 5).

3 Focus on the core's full potential

Thirty years of portfolio expansion – from clothes and watches, to books and amusement parks – have led to Lego neglecting its core business – "creative building" with Lego bricks. When competitors entered the market with cheaper products, the company had to handle a set of internal issues ranging from supply chain complexity to the high product costs in order to regain momentum in the core business. This example shows that if a company's core business is sustainable, like Lego's "creative building", and there are untapped opportunities, the next step should be to focus on and realize the full potential. The two most critical questions here are:

What initiatives are needed to reach strategic full potential?
What initiatives are needed to reach operational full potential?

Initiatives to reach strategic full potential enhance and defend the existing core business. Lego improved its core by tightening its product range as well as through clear customer focus. The latter was achieved by, for example, involving customers in an early stage of the product innovation activities. The goal of reaching the operational full potential is aimed at extracting more from the existing core business. In this regard, activities could include reducing complexity, spinning off non-core activities, and reorganizing. Lego also reorganized its supply chain, divested non-core segments such as the amusement parks, flattened its hierarchy levels, and increased the tempo of its product development and commercialization processes.

4 Expand into close adjacencies

If the core business is sustainable and reached, or is close to its full potential, the right strategic path to follow is to expand the core into adjacent markets (see Zook and Allen, 2003). Vodafone and Nike have successfully managed adjacency growth in the past. Their activities will be discussed in detail below.

The following questions guide managers through this phase:

What is the long list of adjacency options?
Which adjacencies are the most promising? Which should be started/stopped/continued?
What is the repeatable formula behind expansion?

An adjacency map, as shown in Figure 4, is a valuable tool to guide the process of option identification in order to generate a long list. After the full map has been drawn, the relevant growth dimensions need to be identified first. For Nike, the relevant dimensions included geography, footwear, apparel, and equipment. Second, a rigorous screening process needs to identify the specific growth options along these dimensions. Nike, for example, identified different sports as the options along the growth dimension footwear.

However, the individual growth dimensions have individual challenges that must be considered. A potential move is along one's specific value chain. By integrating the value chain forwards or backwards, changes in an industry's profit pool can be mirrored in a company's strategy, allowing participation in an industry's future honey pots. Nevertheless, this is one of the most difficult moves; its challenges include competitive threats and insufficient capabilities to enter the next or prior value chain steps. A successful example of value chain integration is De Beers's move from a diamond wholesaler to a diamond retailer. Other options are the offering of new or modified products and services, selling the existing offering through new distribution channels, to new customer segments, and in new geographic markets. Vodafone's expansion from its UK home turf into Continental Europe and beyond followed this growth path. With a rapid and repeatable screening process of target cellular phone service companies, Vodafone managed to obtain the leading players in the individual markets and gain scale rapidly. The last, and rarest move, is a "white space" move, which is based on a core capability inherent in a company's daily business activities. It is critical to note that the organizations we use as examples selected one growth dimension and a close-to-the-core option for their growth strategy, avoiding moves into unknown terrain.

After completing the long list of adjacency growth options, the selection of the most promising options is the next step. In this regard, assessing previous successful and failed adjacency moves is helpful to understand the success factors. Furthermore, the selection should be based on two criteria: the profit pool attractiveness – based on growth, size, and margins – and the company's ability to win against its competitors in the sector. The ability to win in growth strategies is strongly dependent on the adjacency's relatedness to the core business, as shown in the abovementioned close-to-the-core moves. Therefore, an ex ante relatedness assessment is crucial. Examples that illustrate the risks of too distant growth activities are Mattel, which entered the learning software business, and Anheuser-Busch, which entered the snack business. In both cases, the companies could not rely on their existing capabilities, market knowledge, and customer relationships to foster their growth attempt. Consequently, Mattel had to write off the acquisition price paid for the software company Learning Company and Anheuser-Busch left the snack market due to its repeated losses. In Figure 3, the five dimensions of relatedness are summarized and the decreasing advantages of more distant adjacencies are shown: while the odds of a one-step adjacency being successful are 35%, they decrease to only 8% for three-step adjacency moves. To measure the distance from the core, the dominant areas of overlap – customers, cost, capabilities, and competition – need to be assessed. The less the given overlap and the greater the number of dimensions involved, for example, introducing a new product into a new geographic market, the further the adjacency is from the current core business.

Relatedness to core business	Core	1-step adjacency	2-step adjacency	3-step adjacency	Diversifi-cation
Shared customers	●	●	◑	○	○
Shared costs	●	◕	◑	◑	○
Shared capabilities	●	◕	◑	◑	○
Shared competition	●	◑	◑	○	○
Odds of success	N/A	**35%**	15%	8%	<5%

Economic distance from core
- Less competitive advantage
- Less potential profitability
- Lower share and customer loyalty

Figure 3: Relatedness assessment framework (Source: Bain & Company)

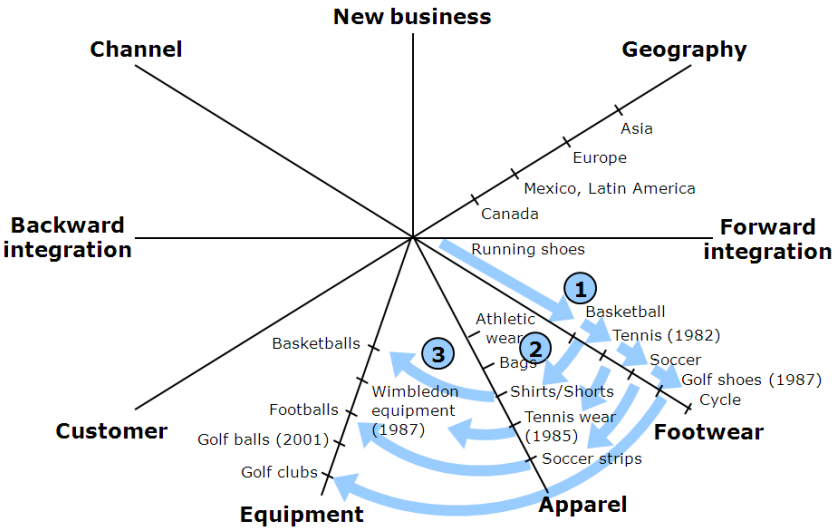


Figure 4: Nike's adjacency map (Source: Bain & Company)

The development of a repeatable formula is the step that separates a successful growth strategist from those who are less successful. By means of this formula, the core assets and capabilities can be repeatedly leveraged for future adjacency moves. A well-known repeatable formula is Nike's entry path into new sports' product segments. Each time when moving into such a segment – for example, basketball, tennis, soccer, and golf – the organization moved from footwear to apparel to equipment, thereafter entering another new sport by following the same path again. This formula has allowed Nike to leverage its experience, increase the speed of its

growth initiatives, to simultaneously fuel its international growth ambitions, and to overtake Reebok and become the global market leader in sportswear. This repeatable formula, together with effective marketing campaigns with leading sports stars – typically the number 1 in every sport discipline – ultimately became a competitive, inimitable competitive resource for Nike.

5 Redefine a new core

If the initial assessment identifies an unsustainable core business, for example, one driven by radical changes in the company's market environment, redefining a new core business might be required (see Zook, 2007). Examples such as IBM, Apple, and De Beers show how redefining the hidden assets inherent in the company led to successful growth initiatives. The following four questions help manage the definition of a new core business.

What is the long list of options for the new core?
Are there hidden assets that can be leveraged into a new core business?
What does the future profit pool look like?
Which options should be pursued in future?

A rigorous screening of opportunities is necessary concerning the adjacency options. The most successful redefinitions build on existing assets or capabilities – often "hidden assets" – and a screening process should therefore take place along three dimensions: customer assets, undervalued growth platforms, and underutilized capabilities. Customer assets include unrecognized segments with leadership potential, privileged access, and underutilized data and information. De Beers, for example, realized that its supply-driven model was losing robustness, which led to a decline in profit. After thoroughly searching for hidden assets in the customer dimension, the transition to the diamond retail business was initiated, which led to a nine-fold value increase over seven years. Undervalued growth platforms such as support organizations for the core, non-core-related businesses with future market potential, and orphan products in dynamic markets are other sources of hidden assets. IBM's global service offering, once a small services unit, was identified as the future core business due to its market potential. It became the focus of attention and investment, and ultimately overtook the hardware business in terms of sales. Finally, underutilized capabilities can arise from underleveraged core capabilities, hidden corporate capabilities, as well as non-core capabilities. Apple found its hidden assets in this dimension. The product design and brand assets were leveraged to create iPod's success and the evolution of a new core business as a music and content distributor with the iTunes platform.

The success of a redefinition strategy based on a company's hidden assets is further supported by a thorough assessment of the industry's future evolvement. The industry's future value chain, its interlinkages, and positioning in the broader ecosystem should be assessed on a scenario basis, including the potential outcomes in the competitive landscape. The selection of the identified long list of activities should therefore be guided by two criteria: the company's ability to win – with its

hidden assets – in the chosen segment, and the segment's attractiveness as assessed in the profit pool analysis.

6 How to become a sustained value creator

The core business's importance in the process of becoming a sustained value creator has been described above. Bain research shows that more than 90% of sustained value creators have a leadership position in their core business. Therefore, identifying the company's core and driving it towards its full potential is the most important growth imperative.

Figure 5 summarizes the three paths that lead to a sustained value creator position around a strong core business. When industries evolve, these steps usually have a consistent cycle and, as shown in the figure, each step requires a focus on its success factors.

In practice, management often falls into the trap of entering new markets with a perceived higher attractiveness than those of its current core business. But if these attractive markets are too distant from the core, and if there is no clear path towards a leadership position, the probability of success is very low. These companies are better advised to focus on the remaining growth opportunities in their core business or close adjacencies in order to fully capitalize on their position in their core market. The rewards of a strong focus on the core business are compelling, as examples such as Lego and Nike demonstrate.

Figure 5: Business evolution cycle (Source: Bain & Company)

References

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management,* 17(1): 99-120.

Zook, C. (2007). Finding Your Next Core Business. *Harvard Business Review*, April 2007.

Zook, C. & Allen, J. (2003). Growth Outside the Core. *Harvard Business Review*, December 2003.

How sustainable are the growth strategies of sustainability entrepreneurs?

Liudmila Nazarkina

Abstract:

Sustainability-driven firms, or firms in the sustainable lifestyle industry producing environmentally friendly and socially just goods and services, pursue a range of growth strategies, including organic growth, acquisitive growth, equity sale to strategic investors, strategic alliances, franchising, and licensing. Each of these growth options may be socially, environmentally, and financially sustainable or unsustainable, depending on the growth strategy arrangements, such as the source and type of funding, as well as on the organizational structure and control. Sustainability entrepreneurs, as well as conventional entrepreneurs seeking socially, environmentally, and financially sustainable growth for their firms, should compare the specific organizational and financial arrangements of growth strategy alternatives and their wider sustainability implications.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_8,

Table of Contents

1 Introduction

"Business is slowing down. The entire world is doing just that, and the adage 'grow or die' is an old manager's tale. [...] Business thus cannot accept or afford anything but a growth pattern that keeps a firm's increase in output and productive capacity in line with increases in its external environment – in other words, sustainable growth" (Kefalas, 1979, p. 40). This was written at the end of 1970s. Interestingly, not much has changed since then. Despite the gloomy prediction of business growth limited by the environment – that is, the availability of natural resources, environmental pollution, and shrinking of the population in the West – the *grow or die* dilemma is far from an old manager's tale. The opposite is true. The levels of production and consumption are growing exponentially. We are using resources 30% faster than the planet can cope with. If the same growth rate continues, we will need two planets by 2050 to maintain our current lifestyles (WWF, 2008). Paradoxically, increased consumption does not make us happier (Jackson, 2008).

Furthermore, despite the hopes that progressive economic growth will improve the well-being of the least well-off, we are still far from reaching the goal of eradicating poverty, reducing hunger, improving access to health and education, and ensuring environmental sustainability everywhere in the world (Millennium Development Goals, 2010). Again, the opposite is true. The globalization of business has resulted in the transfer of negative environmental impacts from the prosperous West to the developing East and South. Media reports abound with stories of labor issues in multinational corporations' international supply chains, including the use of child labor, inhumanely long working hours, and inappropriate working conditions.

Limiting our consumption is only part of the solution to these problems. A family of four probably does not need four cars; we do not necessarily need to own five pairs of jeans, and in winter we do not necessarily need to eat strawberries flown in from halfway across the world[1]. However, we still need some of these products and, to a certain degree, the production-consumption economic cycle is necessary to lift people out of poverty. The question is how these goods are produced, or rather, how we can satisfy our needs for the freedom of movement, quality food and clothing, comfortable accommodation, and energy in ways that minimize negative environmental and social consequences. We may not need four cars to be mobile, but we may need a developed public transportation system, car sharing programs, safe bicycle lanes; we may not need five pairs of jeans that develop holes after a few months, but we may need one or two pairs that last for five years and are preferably made of organic and fair trade cotton; we may not need to eat strawberries in winter, but we can still enjoy strawberries in summer grown by local farmers, even if it means paying a little more for them. This is where we come to the notion of sustainable lifestyles and sustainable entrepreneurship.

[1] Russo (2010) summarizes findings from prior research on the consumption economy which suggests that a typical American woman owns 19 pairs of shoes, but wears only one-quarter of them; Australian households spend over 1,000 AUSD on goods that go to waste; obesity rates in developed countries, especially the US and UK, are very high; however, the percentage of "very happy people" in these countries is lower than in the 1960s.

A sustainable lifestyle can be defined as "patterns of action and consumption, used by people to affiliate and differentiate themselves from others, which: meet basic needs, provide a better quality of life, minimise the use of natural resources and emissions of waste and pollutants over the lifecycle, and do not jeopardise the needs of future generations" (CSfD, 2004, p. 4). Using the same logic of resource preservation and care for the needs of future generations, sustainable entrepreneurship can be defined as an entrepreneurial activity "focused on the preservation of nature, life support, and community in the pursuit of perceived opportunities to bring into existence future products, processes, and services for gain, where gain is broadly construed to include economic and non-economic gains to individuals, the economy, and society" (Shepherd and Patzelt, 2011, p. 137). The sustainable lifestyle industry, also termed LOHAS (lifestyle of health and sustainability), is growing rapidly. Available data suggest that the LOHAS industry in the US grew by 36% in 2008 compared to 2005 – annualized growth of more than 16% – reaching almost 300 billion USD in 2008 (Natural Marketing Institute, 2010). A similar rapid market growth trend is recorded for the UK-based sustainable lifestyle industry, which has grown from 13.5 billion GBP (21 billion USD) in 1999 to 36 billion GBP (56 billion USD) in 2008, an almost threefold increase (Co-operative Bank, 2009).

In academic and practitioner-focused literature, firms in the LOHAS industry are referred to as "sustainability-driven ventures" (Shlange, 2009; Parrish and Foxon, 2009), "sustainability-driven firms" (Parrish, 2010), "mission-driven companies" (Russo, 2003; Russo, 2010), organizations with an "ethical corporate identity" (Balmer et al., 2007), "hybrid organizations" (Boyd et al., 2009), "authentic companies" (Croft, 2005), "companies with a conscience" (Scott and Rothman, 1992), and "social enterprises" (Tracey and Phillips, 2007). The most high-profile examples of such firms are Ben & Jerry's, an ice cream producer that supports a range of social causes, such as organic agriculture and the protection of rainforests; The Body Shop, an organic body care product retailer that supports a wide range of social and environmental programs, such as combating animal testing; Tom's of Maine, which offers organic toothpaste; Stonyfield Farm Yogurt, selling organic and probiotic dairy products; Green & Black's with its organic chocolate; Innocent Drinks, which sells smoothies made of all-natural ingredients; and howies, a manufacturer of organic and durable sports clothing. One would expect sustainability-driven firms to take the question of their growth strategies' sustainability – in terms of financial, social, and environmental sustainability – seriously. How sustainable are the growth strategies of sustainability entrepreneurs?

2 Staying small versus growing big

Let's return to the *grow or die* dilemma. Considering all the environmental and social issues related to growth, would it make more sense for sustainability entrepreneurs to keep their businesses small? Previous research suggests that small firms experience the "liability of smallness" and the "liability of newness" (Stinchcombe, 1965; Freeman et al., 1983; Bruderl and Schussler, 1990; Bruderl and Preisendor-

fer, 1998). Such liabilities impact a small firm's ability to establish long-term relationships with its customers, suppliers, and financiers, thus reducing the firm's chances to survive and grow. In this research tradition, it is assumed that all firms have natural incentives to grow in order to overcome the liabilities of newness and smallness and to achieve economies of scale (Penrose, 1959).

Arguing that this view is overly simplistic, Davidsson (1989) shows that there are many businesses with a low "willingness to grow". These firms prefer to stay small or to grow organically in order to preserve the creative atmosphere of a small firm, maintain full control of the firm, and ensure employee well-being. Davidsson's arguments were further explored by a number of other researchers (e.g., Wiklund et al., 2003; Cliff, 1998; Gimeno et al., 1997), who also provide evidence that an entrepreneurial attitude towards growth – that is, beliefs about the consequences of "more adventurous" growth strategies, as well as beliefs in an entrepreneur's ability to manage rapid growth – may influence the decision to grow slowly or not to grow at all.

However, growing and staying small are not necessarily mutually exclusive, but depend on the criteria used to measure growth. Previous researchers have suggested various growth measurement criteria, including the increase in the number of employees, sales, revenues, market share, customers, dividends, assets, etc. (Weinezimmer et al., 1998). Therefore, it is possible to imagine a scenario of a firm increasing its sales, thus growing, by maintaining the same number of employees, thus not growing, or staying small. Using an example from the sustainable lifestyle industry, it has been shown that the growth of Cafédirect, a fair trade coffee distributor in the UK, was achieved through extensive use of networks, while the number of employees has never exceeded 35 people (Davies et al., 2010). Additionally there is some evidence that academics and entrepreneurs view growth differently; for entrepreneurs, growth is very much about organizational development, rather than increases in the number of employees or sales (Achtenhagen et al., 2010).

However, in some industries, scale may be crucial in order to survive and successfully compete with market incumbents. For sustainability-driven firms, increasing scale also means bringing their sustainability innovation to more people, thus contributing to their mission of transforming the world economy into a more sustainable one. Fair trade researchers suggest that there is a certain threshold in fair trade companies' turnover that is very difficult to surpass (Randall, 2005). Some even quantify it as a 3% market share (e.g., Davies et al., 2010). Fair trade companies' limited growth may be partially attributed to their use of a focus (niche) strategy aimed at increasing the number of specialized retail outlets, rather than at increasing penetration in mainstream outlets. "Going mainstream" is proposed as a more appropriate growth strategy for fair trade companies (Randall, 2005).

3 Going mainstream

What options do sustainability-driven firms have to "go mainstream" in order to increase the distribution of their sustainable goods and services? The following quote from Bo Burlingham's (2005) book *Small Giants: Companies That Choose to*

Be Great Instead of Big is an excellent introduction to the subsequent discussion: "Up to that point, Erickson had been convinced that he had no real choice but to sell. His two largest competitors – Power Bar and Balance Bar – had recently been sold to Nestlé and Kraft, respectively. He and Thomas, who was the CEO, were terrified of competing head-to-head against multibillion-dollar conglomerates that had the financial resources to wipe them out overnight. They believed that, by selling to another giant company, they could protect Clif Bar and its employees, since they themselves would remain in control" (p.7).

3.1 Selling firm ownership

It appears that, as a strategy to go mainstream, sustainability entrepreneurs are increasingly choosing to sell their firms to larger mainstream competitors. Ben & Jerry's, The Body Shop, Tom's of Maine, Stonyfield Farm Yogurt, Green & Black's, Burt's Bees, Innocent Drinks, and Honest Tea are only some of the examples of sustainability-driven firms sold to large corporations – Unilever, L'Oréal, Colgate, Danone, Cadbury Schweppes, Clorox, and Coca-Cola, respectively. As suggested by Innocent Drinks, "having Coke as an investor brings big advantages: we can leverage Coke's route to markets in countries we wouldn't otherwise be able to operate in, and they help us access better rates on everything from media to oranges. [...] This deal will help us get our little bottles of healthiness to many more people" (Financial Times, 2010).

In the management and entrepreneurship literature, the seller's perspective on mergers and acquisitions remains relatively unexplored. A recent literature stream, initiated by Graebner and Eisenhardt (2004), suggests that small entrepreneurial firms may be "pushed" towards a sell-out when they face difficult strategic hurdles (e.g., raising finance or hiring a new CEO). Alternatively, they may be "pulled" towards a sell-out by the organizational support offered by an acquirer (e.g., an opportunity to use a parent company's production facilities or distribution channels), as was the case with Innocent Drinks. Entrepreneurs may also be looking to sell their firm when they try to exit the business as a result of either business failure or business success (Wennberg et al., 2010).

Although firm sale may result in a lucrative reward for a business founder and, potentially, in performance improvement for a sold firm, there is also extensive evidence that sell-outs can lead to significant organizational changes in the acquired firm, including employee layoffs. The initial layoffs can have an influence on the remaining employees' morale, which increases the "brain drain" and impacts the firm productivity and innovative capabilities (Krug and Hegarty, 2001; Paruchuri et al., 2006; Kapoor and Lim, 2007; Spedale et al., 2007). This is exactly what happened to Ben & Jerry's following its sale to Unilever in 2000. Shortly after the acquisition, during which time organizational changes were prohibited, Unilever undertook significant layoffs and a full integration of the manufacturing functions. Only Ben & Jerry's marketing autonomy was maintained (Mirvis, 2008), although the initial promise had been to retain the company's manufacturing base and to continue supporting a number of social programs, such as buying milk from farmers who did not

use growth hormones, and purchasing from fair trade suppliers (Independent, 2008). In the 2004 *Social and Environmental Report*, Ben & Jerry's sadly admitted that "we are beginning to look like the rest of corporate America" (Ben & Jerry's, 2004, p. 29).

A somewhat similar story happened to howies, which was founded by David and Clare Hieatt, two sustainability entrepreneurs, in 1995. Their mission was to "make people think about the world we live in" by offering fashionable and environmentally friendly clothing. Since its creation in 1995, the business grew rapidly. For more than 10 years, David and Clare Hieatt financed howies' expansion by investment from family and friends. However, in 2006, howies was sold to Timberland, a large sports clothing manufacturer and retailer. "We knew we couldn't give howies any more money. We didn't have any more to give. [...] So the answer for both of us was clear and that was to find a like-minded company to help howies become brilliant at being howies" (howies blog, 2006). Later on in this interview, David Hieatt asserted that howies would "maintain creative control of howies as we always have done, and the same management team will continue to run the company from good old Cardigan Bay. [...] So really it's business as usual" (howies blog, 2006). However, in October 2009, three years after the sale of howies to Timberland, David Hieatt announced that he was leaving the company: "Yesterday was my last day at howies. It was my choice to leave. It was a hard thing to do and easy thing to do. The moral of the story is simple. If you find something you love, you should never sell your love. When other people own your dream, destiny is no longer in your hands" (howies blog, 2009).

However, not all acquisitions are equally traumatic. Haspeslagh and Jemison (1991) distinguish between three types of acquisitions: (1) preservative acquisition, which means low interdependence between an acquirer and an acquired firm, and high autonomy on the part of the acquired firm; (2) symbiotic acquisition, which means high interdependence, but also a high degree of autonomy on the part of the acquired firm; and (3) absorptive acquisition, which means full integration, with limited autonomy on the part of the acquired firm. Previous research suggests that a certain degree of integration is required in order to exploit the synergies between an acquirer and an acquired firm (Schweizer, 2005). Nevertheless, if the acquired firm has greater autonomy, this also leads to improved performance by both the acquirer (Datta and Grant, 1990; Very and Lubatkin, 1997) and the acquired firm (Ranft and Lord, 2002; Kapoor and Lim, 2007; Paruchuri et al. 2006). Thus, Mirvis (2008) suggests that the preservative and symbiotic acquisitions of Tom's of Maine, Green & Black's, and Stonyfield Farm Yogurt were effective in retaining the innovative capacity of sustainability-driven firms, while the absorptive acquisition of Ben & Jerry's resulted in the loss of innovativeness, low employee morale, and product quality problems.

It seems that sustainability-driven firms and large corporations have also been learning from the pioneers' mistakes and are structuring acquisition deals better. For example, in 2001, Coca-Cola purchased 100% ownership in Odwalla, a US natural juice producer, with Odwalla experiencing the full consequences of the 2008 boycott of Coca-Cola products. The Innocent Drinks/Coca-Cola and Honest Tea/Coca-Cola deals were structured differently. In 2008, Coca-Cola purchased a minority stake (18%) in Innocent Drinks, followed by the purchase of another 40%

in 2010. Similarly, in 2008, Coca-Cola acquired a 40% stake in Honest Tea, a US-based organic bottled tea company; it acquired the remaining 60% stake in March 2011. The interaction between the acquired sustainability-driven firms and Coca-Cola seems productive, and both Innocent Drinks and Honest Tea are growing steadily, reaching new markets, and new consumers. For example, since receiving its first investment in 2008, Honest Tea has increased its distribution from about 15,000 outlets in 2008 to more than 75,000 in 2011. Furthermore, it has introduced a ground-breaking Honest Tea plastic bottle that uses 22% less material; it has also doubled its product line (Business Wire, 2011). Commenting on the deal, Honest Tea's co-founder, Seth Goldman, noted: "We've been engaged to Coke for the past three years, and now we are getting married. When you're engaged to somebody, you're around them enough that you can gauge any bad habits. So we don't really expect any surprises, and we certainly have developed a good working relationship" (Washington Post, 2011).

Let's pause and think about the financial, environmental, and social sustainability of growth by selling firm ownership to a large corporation. First, this appears to be a fairly risky strategy. As Katila et al. (2008) note, entrepreneurial firms selling equity to strategic investors often have to "swim with sharks" (p. 295). Even when operational control or "creative control" of the firm is retained, sell-outs do not always lead to firm growth and, sometimes, result in quite the opposite: Destruction of the business model, mission loss, and even the company's demise. The latter occurred in the sell-out of PJ Smoothies, a UK smoothies market pioneer, which saw the discontinuation of production in 2009, four years after its sale to PepsiCo. Unsuccessful sell-outs may also have significant implications for the firm stakeholders, such as employee layoffs or cancellation of long-term supplier contracts, as well as indirect environmental consequences, such as the parent firm's limited commitment to environmental protection and resource preservation, which have repercussions for the acquired firm's production processes.

However, there are reasons to believe that large corporations are not purposefully destroying acquired firms. In the cases of Innocent Drinks, Honest Tea, and Coca-Cola, properly structured acquisition deals may be equally beneficial for both sides. Furthermore, such deals may form part of sustainability-driven firms' missions. In a recent publication, Hockerts and Wuestenhagen (2010) refer to sustainability-driven firms as "emerging Davids," as opposed to "greening Goliaths" or large corporations. Emerging Davids have a pioneering role: They create new market niches, slowly grow them, and eventually, often unwillingly, attract attention from greening Goliaths, who see commercial opportunities for entering new market niches. Ultimately, the interaction between Davids and Goliaths leads to progressive market transformation towards a more environmentally and socially sustainable economy. Although Coca-Cola has not yet discontinued the production of its flagship Coca-Cola drink, which has been extensively criticized for the environmental impact of water-intensive production in drought-prone areas, it is now launching its own sustainable lifestyle products. These products are Cascal, a handcrafted, all-natural soda, and Sokenbicha, unsweetened bottled teas made of "natural botanicals." Interestingly, Coca-Cola positions them as stand-alone brands, sim-

ilar to Innocent Drinks and Honest Tea, and customers cannot identify the true owner by looking at the product label, its Facebook page, or Twitter stream. Can Coca-Cola still be Coca-Cola if it no longer produces any Coca-Cola? The answer is yes. Corporate history contains many examples of corporations moving into completely new industries, while retaining their names (e.g., General Electric, AT&T, Nokia, and Procter & Gamble).

Despite numerous sell-out examples, selling a sustainability-driven firm to a large corporation is not the only possible strategy to achieve scale and go mainstream. Sustainability entrepreneurs can establish strategic alliances with larger firms, franchise their businesses, license their products and technologies, or merge with or acquire other small sustainability-driven firms. Let's take a look at these growth options and their sustainability implications.

3.2 Acquisitive growth

Acquisitive growth is exactly the opposite of a sell-out. This strategy is often pursued by firms seeking to increase the scale of operations, to acquire new capabilities and resources while entering new geographic markets (Wiklund and Shepherd, 2009), and to revitalize a firm, particularly one which has been growing organically over long periods (Penrose, 1959; Vermeulen and Barkema, 2001). An example of acquisitive growth in the sustainable lifestyle industry is Zipcar, a US-based car-sharing company popular among college students and city residents. Zipcar allows customers to rent cars at an hourly or daily rate, and park in convenient spots in the city centre where parking is scarce and expensive. Despite persistent annual losses, amounting to an accumulated deficit of 65 million USD in 2010, Zipcar has expanded rapidly, both in the US and internationally, by acquiring Flexcar (US) in 2007, Avancar (Spain) in 2009 (minority acquisition), and Streetcar (UK) in 2010. After receiving six rounds of venture capital (VC) finance, Zipcar raised 174 million USD in the initial public offering (IPO) on the NASDAQ Stock Exchange in April 2011, 30% more than initially planned (Wall Street Journal 2011). Although Zipcar has warned investors that losses should be expected for 2011-12, as it is still making many upfront investments in its car fleet, there is currently much excitement about Zipcar shares.

How did Zipcar manage to achieve scale without selling out to a large corporation? Similar to Innocent Drinks and howies, Zipcar faced strong competition from market incumbents such as Hertz Global Holdings Inc, which also started rival car-sharing services. Zipcar also faced increasing competition from the next generation of car-sharing start-ups such as RelayRides, GetAround, and Spride Share. These start-ups were founded on the idea of sharing personal vehicles, owned by people in the neighborhood, instead of by a business entity like Zipcar. A lack of financing is also not sufficient to explain the difference in growth strategy choices. After all, howies made a conscious decision not to attract VC financing, as it "would put too much pressure on us to grow and, maybe, force us to start cutting corners that we would have found hard to live with" (howies blog, 2006). VC financing is also no panacea against a sell-out, since VC investors seek exit strategies from their invest-

ments. If an IPO is not an option, the next-best exit strategy is a trade sale to a large corporation. This is what happened to, for example, VC-backed Burt's Bees, a natural personal care products company, which was sold to Clorox in 2004, and to VC-backed Kettle Foods, a producer of hand-cooked natural crisps, which was sold to Diamond Foods in 2010.

Finally, how sustainable is growth by means of acquiring other firms? Although acquisitive growth creates the illusion of independent growth, it is mostly accompanied by partial loss of ownership. This loss is due to the company requiring external financing, either in the form of equity sale to private equity (PE) investors, or an IPO. Unfortunately, PE ownership or listing on a stock exchange can lead to a reduction in the spread of "allowable mission-drivenness" (Russo 2010, p. 176) owing to fiduciary and reporting constraints. PE investor or firm shareholders may therefore enforce their fiduciary prerogatives and limit the sustainability-driven firm's social and environmental activities. Acquisitive growth is also accompanied by post-acquisition integration challenges faced by the acquirer. Previous research suggests that the failure to achieve the correct level of integration between an acquirer and an acquired firm (over-integration or under-integration) can lead to value destruction on both sides (Pablo, 1994; Schweizer, 2005; Graebner et al., 2010).

We have investigated examples of firm growth financed by strategic investors (sell-outs to large corporations) and acquisitive growth financed by VC/PE investors or an IPO; however, firms can grow in various other ways that do not necessarily require financing from strategic or financial investors. These forms of company growth can be grouped under "hybrid growth."

3.3 Hybrid growth

Hybrid growth strategies involve growth via the use of long-term contractual organizational arrangements with both parties maintaining autonomy, but with the relationship controlled by a special adaptive mechanism (Williamson, 1991; Shane, 1996). These strategies include franchising, licensing, and strategic alliances (McKelvie and Wiklund, 2010). Franchising can be defined as "a legal contract between a trademark owner (franchisor) and a local user (franchisee) to sell products or services under the owner's trademark employing a production process developed by the franchisor" (Michael, 1996, p. 59). The main advantage of franchising is rapid expansion that can lead to economies of scale with a relatively modest investment by a franchisor, since franchisees invest their own capital (Combs and Ketchen, 1999). In addition to bringing in additional capital and a new revenue stream, franchising is especially beneficial for firms pursuing an internationalization strategy, since franchisees also introduce competent new staff with knowledge of local markets (Combs and Ketchen, 2003).

Licensing, defined as selling the rights to an intellectual property (technology or product) to a licensee (McKelvie and Wiklund, 2010), is another strategy to raise an additional revenue stream and, thus, grow a firm. Large corporations, such as IBM and Dow Chemicals, generate hundreds of millions of USD through their licensing activities (Lichtenthaler, 2008). A specific example of technology licensing is

Motorola's decision to sell property rights to its GMS technology to direct competitors Nokia and Ericsson (Kline, 2003). In addition to generating a new revenue stream, licensing may be a strategic option allowing small firms to overcome deficiencies in their manufacturing base and, thus, commercialize their innovations even if they cannot launch manufacturing independently (Fosfuri, 2006). Despite their benefits, both franchising and licensing may lead to opportunistic behavior by franchisees (e.g., reduced product quality for the sake of increased profitability) and licensees (e.g., leaking proprietary information), which may have possible reputational effects (McKelvie and Wiklund, 2010).

Finally, entrepreneurial firms may establish strategic alliances with similar-sized firms, or with larger and more experienced firms. A strategic alliance can be defined as an "administrative arrangement to govern an incomplete contract between separate firms in which each partner has limited control" (Gomes-Casseres 1997, p. 34). Strategic alliances can take the form of joint ventures, cooperative marketing, or R&D arrangements. Joint ventures are equity-based forms, while other forms of strategic alliances are contract based. The main reason for entering into strategic alliances is to gain access to new capabilities and resources (Das and Teng, 2000), particularly when entering new geographic markets (Hollenstein 2005; Kirby and Kaiser, 2003; Chen and Huang, 2004). A strategic alliance may therefore be a less risky and a less costly internationalization strategy than organic or acquisitive growth. Strategic alliances are also particularly attractive for small firms, since they increase their survival rates (Baum and Oliver, 1991) and improve their competitive positions (Beekman and Robinson, 2004). Additionally, strategic alliances provide small firms with an opportunity to "become big" by "staying small" and to compete with larger companies as a constellation of firms, rather than on their own (Gomes-Casseres, 1997). However, strategic alliances can have the same range of opportunism problems as franchising and licensing.

The cited case of Cafédirect is a good example of hybrid growth in the sustainable lifestyle industry. Most of Cafédirect's growth can be attributed to its increased networking activities through partnerships with other organizations (Davies et al., 2010). The benefits and risks of franchising, pursued by sustainability-driven ventures, have been analyzed in the context of Aspire's social venture franchising system. As identified by Tracey and Jarvis (2007), access to financial resources, managerial talent, and local knowledge were the main factors that motivated Aspire to set up a franchising system. However, in the case of Aspire, the franchising experiment resulted in the collapse of the whole system, since (1) the franchisor did not have the necessary experience to evaluate the potential franchisees' organizational competences to achieve economic and social goals; (2) the franchisor and franchisees' goals were not fully aligned (economic viability vs. social service provision); and (3) efficient franchisee monitoring was not put in place.

Unfortunately, franchising and licensing are underexplored in the academic literature on firm growth, including the literature on sustainability-driven firms. With respect to strategic alliances, there exists a large body of knowledge on the tradeoffs between alliances and acquisitions (e.g., Villalonga and McGahan 2005). However, there is only anecdotal evidence on the use of strategic alliances by sustainability-

driven firms. Despite limited knowledge of the prevalence of hybrid strategies among sustainability entrepreneurs, one may hypothesize that these forms of growth are less risky and potentially more socially, environmentally, and financially sustainable, since there is a smaller likelihood of negative consequences of "swimming with sharks" strategies or risks involved in collaborating with VC/PE financiers in acquisitive growth.

4 Recommendations for sustainability entrepreneurs

What is therefore the most sustainable growth strategy for sustainability entrepreneurs? There is probably no right or wrong strategy (Table 1). To cite Parrish (2010), "two sets of entrepreneurs, those from Chumbe Island and Hill Holt Wood, decided they could be most effective environmentally and socially by stopping growth at a certain level, while the other two, those from *Native*Energy and ForesTrade, decided they could be most effective by growing their enterprises indefinitely. Taking one example, the case of ForesTrade, the entrepreneurs felt that the more producers they could partner with as suppliers, the greater their positive social and environmental impacts would be" (p. 519). It would therefore be incorrect to suggest that slow growth, or even no growth, is the best strategy for all sustainability entrepreneurs. For some, continuous growth is the only way to strongly contribute to the sustainable development of the systems of which they are part. The question is: Exactly how is growth achieved?

If the chosen growth strategy leads to high-risk situations fraught with negative social and environmental consequences, it is probably not the most sustainable way to grow a sustainability-driven firm. Straight sell-outs seem to be the most risky strategy, with a high likelihood of undesirable consequences, while two-stage sell-outs – which allow for testing the relationship between the two firms prior to committing to a full ownership change – may be a more sustainable way that a firm, which chooses to use a strategic investor's help to increase scale, can grow. Similarly, hybrid growth strategies, such as franchising, may be a sustainable way to grow – if proper coordination mechanisms are put into place and a franchisor and franchisee's goals are aligned. However, it may also be an unsustainable growth strategy, which may lead to a firm's demise if the hazards inherent in hybrid growth arrangements are not considered. Finally, acquisitive growth may or may not be sustainable either; everything depends on the alignment of the goals and values between a sustainability-driven firm and its VC/PE investors. For example, so-called social venture capital firms (e.g., Good Capital) allow sustainability-driven firms in their portfolio to pay off VC investment by raising the funds internally or through debt financing, instead of selling their portfolio companies to a large corporation (Russo, 2010). There is also much talk about a "social stock exchange" that would allow external finance provision for sustainability-driven firms with active expansion strategies. Finally, sustainability-driven firms pursuing acquisitive growth and financing it via listing on a stock exchange may be using "clever" strategies to retain control of the firm. They may, for example, issue two types of stock,

A shares and B shares, with different voting rights: A shares are traded publicly, while B shares, with more voting rights, are owned by the firm management, thus preserving the spread of "mission-drivenness" (Russo, 2010).

To sum up, while sustainability entrepreneurs have many options to grow their firms, each of these growth options may be both sustainable and unsustainable, depending on the growth strategy arrangements, such as the source and type of funding and organizational structure and control. Therefore, the recommendation for sustainability entrepreneurs, as well as for conventional entrepreneurs seeking socially, environmentally, and financially sustainable growth for their firms, would be to think holistically and compare the specific organizational and financial arrangements of growth strategy alternatives and their wider social and environmental sustainability implications.

Table 1: Sustainability implications of firm growth strategies

Growth strategy	Sustainability implications*	
	Positive	**Negative**
Organic growth	• preserving the creative atmosphere of a small firm • maintaining full control over the firm • ensuring employee well-being • limited environmental impact	• limited ability to distribute sustainability innovations
Equity sale to strategic investors	• access to new markets and distribution channels • access to new sources of financing • access to new capabilities • harvest strategy for firm founders • gradual industry transformation by influencing acquirers' sustainability practices	• loss of control over the firm • employee layoffs • reduced firm productivity • cancellation of contracts with long-term suppliers • reduced commitment to environmental protection and resource preservation
Acquisitive growth	• increasing the scale of operations • access to new capabilities • firm revitalization	• high capital requirements leading to partial equity sale • value destruction due to over-integration or under-integration of the acquired firm
Franchising	• rapid expansion • economies of scale • entry into new markets • limited capital requirements • new revenue stream • access to new capabilities	• opportunistic franchisee behavior (e.g., reduced product quality)
Licensing	• new revenue stream • commercialization of innovation	• opportunistic licensee behavior (e.g., leakage of proprietary information)
Strategic alliances	• access to new capabilities • increase in survival rates • increase in competitiveness	• opportunistic behavior by a strategic alliance partner

* Here, sustainability implications include internal sustainability (e.g., business continuity and employee well-being), as well as external sustainability (e.g., implications for external stakeholders, such as local communities and the natural environment). This table presents *potential* sustainability implications; the *actual* sustainability implications will depend on the strategy implementation (i.e. the source and type of funding, as well as the organizational structure and control).

References

Achtenhagen, L., Naldi, L., & Melin, L. (2010). "Business growth"—Do practitioners and scholars really talk about the same thing? *Entrepreneurship Theory and Practice*, 34(2), 289-326.

Balmer, J. M. T., Fukukawa, K., & Gray, E. R. (2007). The nature and management of ethical corporate identity: A commentary on corporate identity, corporate social responsibility and ethics. *Journal of Business Ethics*, 76(1), 7-15.

Baum, J. A. C., & Oliver C. (1991). Institutional linkages and organizational mortality. *Administrative Science Quarterly*, 36, 187-218.

Beekman, A. V., & Robinson, R. B. (2004). Supplier partnerships and the small, high-growth firm: selecting for success. *Journal of Small Business Management*, 42(1), 59-77.

Ben & Jerry's (2004). Social and environmental assessment 2004: One scoop at a time. http://www.benjerry.com/company/sear/2004/sea_2004.pdf. Accessed 6 January 2011.

Boyd, B., Henning, N., Reyna, E., Wang, D. E., & Welch, M. D. (2009). *Hybrid organizations: New business models for environmental leadership.* Sheffield: Greenleaf Publishing.

Burlington, B. (2005). *Small giants: Companies that choose to be great instead of big*. New York: Penguin.

Bruderl, J., & Preisendorfer, P. (1998). Network support and the success of newly founded businesses. *Small Business Economics*, 10(3), 213-225.

Bruderl, J., & Schussler, R. (1990). Organizational mortality: The liabilities of newness and adolescence. *Administrative Science Quarterly*, 35(3), 530-547.

Business Wire (2011). Honest tea joins the Coca-Cola family: Organic bottled tea company well positioned for continued growth. 1 March 2011.

Chen, H. L., & Huang, Y. (2004). The establishment of global marketing strategic alliances by small and medium enterprises. *Small Business Economics*, 22, 365-377.

Cliff, J. E. (1998). Does one size fit all? Exploring the relationship between attitude towards growth, gender, and business size. *Journal of Business Venturing*, 13(6), 523-542.

Combs, J. G., & Ketchen, D. J. (1999). Can capital scarcity help agency theory explain franchising? Revisiting the capital scarcity hypothesis. *Academy of Management Journal*, 42: 196-207.

Combs, J. G., & Ketchen, D. J. (2003). Why do firms use franchising as an entrepreneurial strategy? A meta-analysis. *Journal of Management*, 29: 443-465.

Co-operative Bank (2009). Ten years of ethical consumerism 1999-2008. The Co-operative Bank and Ethical Consumer Research Association (ECRA). http://www.goodwithmoney.co.uk/assets/Ethical-Consumerism-Report-2009.pdf. Accessed 9 May 2011.

Croft, N. (2005). *Authentic business: How to create and run your perfect business*. Chichester: Capstone Publishing Limited.

CSfD (2004). Every little bit helps… Overcoming the challenges to researching, promoting and implementing sustainable lifestyles. Centre for Sustainable Development, University of Westminster. http://home.wmin.ac.uk/cfsd/reports/Sustainable_lifestlyes_summary_report.pdf Accessed 9 May 2011.

Das, T. K., & Teng, B. S. (2000). A resource-based theory of strategic alliances. *Journal of Management*, 26, 31-61.

Datta, D. K., & Grant, J. H. (1990). Relationships between type of acquisition, the autonomy given to the acquired firm, and acquisition success: An empirical analysis. *Journal of Management*, 16(1), 29-44.

Davidsson P. (1989). Entrepreneurship – and after? A study of growth willingness in small firms. *Journal of Business Venturing*, 4(3), 211-226.

Davies, I. A., Doherty, B., & Knox, S. (2010). The rise and stall of a fair trade pioneer: The Cafédirect Story. *Journal of Business Ethics*, 92(1), 127-147.

Financial Times (2010). Judgment Call: How far should companies go in pursuit of revenues? Stefan Stern. 21 April 2010. http://www.ft.com/intl/cms/s/0/67af3842-4cdc-11df-9977-00144feab49a.html#axzz1RujPkH77. Accessed 6 January 2011.

Fosfuri, A. (2006). The licensing dilemma: Understanding the determinants of the rate of technology licensing. *Strategic Management Journal*, 27, 1141-1158.

Freeman, J., Carroll, G. R., & Hannan, M. T. (1983). The liability of newness: Age dependence in organizational death rates. *American Sociological Review*, 48(5), 692-710.

Gimeno, J., Folta, T. B., Cooper, A. C., & Woo, C. Y. (1997). Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, 42, 750-783.

Gomes-Casseres, B. (1997). Alliance strategies of small firms. *Small Business Economics*, 9, 33-44.

Graebner, M. E., & Eisenhardt, K. M. (2004). The seller's side of the story: Acquisition as courtship and governance as syndicate in entrepreneurial firms. *Administrative Science Quarterly*, 49(3), 366-403.

Graebner, M. E., Eisenhardt, K. M., & Roundy, P. T. (2010). Success and failure in technology acquisitions: Lessons for buyers and sellers. *Academy of Management Perspectives*, 24(3), 73-92.

Haspeslagh, P. C., & Jemison, D. B. (1991). *Managing acquisitions: Creating value through corporate renewal*. New York: Free Press.

Hockerts, K., & Wuestenhagen, R. (2010). Greening Goliaths versus emerging Davids — Theorizing about the role of incumbents and new entrants in sustainable entrepreneurship. *Journal of Business Venturing*, 25, 481-492.

Hollenstein, H. (2005). Determinants of international activities: Are SMEs different? *Small Business Economics*, 24, 431-450.

howies blog (2006). Exciting news. http://www.howies.co.uk/content.php?x SecId=56&viewblog=557, 4 December 2006. Accessed 8 June 2011.

howies blog (2009). Stay hungry. Stay foolish. http://brainfood. howies.co.uk/2009/10/11295, 27 October 2009. Accessed 8 June 2011.

Independent (2008). The food chain: How big business bought up the ethical market. Martin Hickman, 14 June 2008. http://www.independent.co.uk/life-style/food-and-drink/features/the-food-chain-how-big-business-bought-up-the-ethical-market-845854.html. Accessed 6 January 2011.

Jackson, T. (2008). The challenge of sustainable lifestyles. In *State of the World 2008: Innovations for a Sustainable Economy*, Chapter 4. Worldwatch Institute.

Kapoor, R., & Lim, K. (2007). The impact of acquisitions on the productivity of inventors at semiconductor firms: A synthesis of knowledge-based and incentive-based perspectives. *Academy of Management Journal*, 50(5), 1133-1155.

Katila, R., Rosenberger, J. D., & Eisenhardt, K. M. (2008). Swimming with sharks: technology ventures and corporate relationships. *Administrative Science Quarterly*, 53(2), 295-332.

Kefalas, A. G. (1979). Toward a sustainable growth strategy. *Business Horizons*, April, 34-40.

Kirby, D. A., & Kaiser, S. (2003). Joint ventures as an internationalisation strategy for SMEs. *Small Business Economics*, 21, 229-242.

Kline, D. (2003). Sharing the corporate crown jewels. *MIT Sloan Management Review*, 44, 89-93.

Krug, J. A., & Hegarty, W. H. (2001). Predicting who stays and leaves after an acquisition: A study of top managers in multinational firms. *Strategic Management Journal*, 22(2), 185-196.

Lichtentaler, U. (2008). Externally commercializing technology assets: An examination of different process stages. *Journal of Business Venturing*, 23, 445-464.

McKelvie, A., & Wiklund, J. (2010). Advancing firm growth research: A focus on growth mode instead of growth rate. *Entrepreneurship Theory and Practice*, 34(2), 261-288.

Millennium Development Goals (2010). 2010 progress chart. http://unstats.un.org/unsd/mdg/Resources/Static/Products/Progress2010/MDG_Report_2010_Progress_Chart_En.pdf. Accessed 2 June 2011.

Michael, S. E. (1996). To franchise or not to franchise: An analysis of decision rights and organizational form shares. *Journal of Business Venturing*, 11(1), 57-71.

Mirvis, P. H. (2008). Can you buy CSR? *California Management Review*, 51(1), 109-116.

Natural Marketing Institute (2010). LOHAS market size: U.S. consumer sales soar to close to $300 billion. LOHAS journal. Spring 2010. http://www.lohas.com/sites/default/files/lohasmarketsize.pdf

Pablo, A. L. (1994). Determinants of acquisition integration level: A decision-making perspective. *Academy of Management Journal*, 37(4), 803-836.

Parrish, B. D. (2010). Sustainability-driven entrepreneurship: Principles of organization design. *Journal of Business Venturing*, 25, 510-523.

Parrish, B. D. & Foxon, T. J. (2009). Sustainability entrepreneurship and equitable transitions to a low-carbon economy. *Greener Management International*, 55, 47-62.

Paruchuri, S., Nerkar, A., & Hambrick, D. C. (2006). Acquisition integration and productivity losses in the technical core: Disruption of inventors in acquired companies. *Organization Science*, 17(5), 545-562.

Penrose, E. T. (1959). *The theory of the growth of the firm*. Reprint. 3rd ed. New York: Oxford University Press.

Randall, D. C. (2005). An exploration of opportunities for the growth of the fair trade market: Three cases of craft organisations. *Journal of Business Ethics*, 56(1), 55-67.

Ranft, A. L., & Lord, M. D. (2002). Acquiring new technologies and capabilities: A grounded model of acquisition implementation. *Organization Science*, 13(4), 420-441.

Russo, M. V. (2003). The emergence of sustainable industries: Building on natural capital. *Strategic Management Journal*, 24(4), 317-331.

Russo, M. V. (2010). *Companies on a mission: Entrepreneurial strategies for growing sustainably, responsibly, and profitably*. California: Stanford University Press.

Schweizer, L. (2005). Organizational integration of acquired biotechnology companies into pharmaceutical companies: The need for a hybrid approach. *Academy of Management Journal*, 48(6), 1051-1074.

Scott, M., & Rothman, H. (1992). *Companies with a conscience*. New York: Birch Lane Press.

Shane, S. A. (1996). Hybrid organizational arrangements and their implications for firm growth and survival: A study of new franchisors. *Academy of Management Journal*, 39, 216-234.

Shepherd, D. A., & Patzelt, H. (2011). The new field of sustainable entrepreneurship: Studying entrepreneurial action linking "what is to be sustained" with "what is to be developed". *Entrepreneurship Theory and Practice*, 35(1), 137-163.

Shlange, L. E. (2009). Stakeholder identification in sustainability entrepreneurship: The role of managerial and organisational cognition. *Greener Management International*, 55, 13-32.

Spedale, S., Van Den Bosch, F. A .J., & Volberda, H. W. (2007). Preservation and dissolution of the target firm's embedded ties in acquisitions. *Organization Studies*, 28(8), 1169-1196.

Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (Eds.), *Handbook of Organizations* (pp. 153-93). Chicago: Rand McNally.

Tracey, P., & Jarvis, O. (2007). Toward a theory of social venture franchising. *Entrepreneurship Theory and Practice*, 31(5), 667-685.

Tracey, P., & Phillips, N. (2007). The distinctive challenge of educating social entrepreneurs: A postscript and rejoinder to the special issue on entrepreneurship education. *Academy of Management Learning & Education*, 6(2), 264-271.

Vermeulen, F., & Barkema, H. (2001). Learning through acquisitions. *Academy of Management Journal*, 44, 457-76.

Very, P., & Lubatkin, M. (1997). Relative standing and the performance of recently acquired European firms. *Strategic Management Journal*, 18(8), 593-614.

Villalonga, B., & McGahan, A. M. (2005). The choice among acquisitions, alliances and divestitures. *Strategic Management Journal*, 26, 1183-1208.

Wall Street Journal (2011). Zipcar IPO soars 66% out of the gate. 14 April 2011.

Washington Post (2011). Coke wraps up deal to buy Honest Tea. 2 March 2011.

Weinezimmer, L. G., Nystrom, P. C., & Freeman, S. J. (1998). Measuring organizational growth: Issues, consequences and guidelines. *Journal of Management*, 24(2), 235-262.

Wennberg, K., Wiklund, J., DeTienne, D. R., & Cardon, M. S. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing*, 25, 361-375.

Wiklund, J., Davidsson, P., & Delmar, F. (2003). What do they think and feel about growth? An expectancy-value approach to small business managers' attitudes toward growth. *Entrepreneurship Theory and Practice*, 27(3), 247-270.

Wiklund, J., & Shepherd, D. A. (2009). The effectiveness of alliances and acquisitions: The role of resource combination activities. *Entrepreneurship Theory and Practice*, 33(1), 193-212.

Williamson, O. E. (1991). Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly*, 36(2), 269-296.

WWF (2008). One planet mobility: A journey towards a sustainable future. WWF-UK. http://assets.wwf.org.uk/downloads/opm_report_final.pdf. Accessed 9 May 2011.

Boards' contribution to organizational growth: Effectiveness as a critical success factor

Michèle F. Rüdisser

Abstract:

As a guardian of corporate values, boards are not only ultimately responsible for organizations' strategic direction, but also play a decisive role concerning balanced growth. However, not all boards place equal emphasis on identifying and populating strategic growth opportunities. Instead of leaving everything to management, board members could make a difference by becoming more actively involved in a company's affairs.

In order for a board to be management's true strategic partner, it is of prime importance that board members show a strong personal commitment to their work. Consequently, it is crucial to make sufficient time available for individual board mandates and to ensure effective working procedures when working as a group. To improve the quality of board involvement and foster organizational growth, each board member could, for instance, be asked to regularly submit innovative ideas for the board's agenda and to vigorously contribute to the list of key strategic decisions that need to be made.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_9,

Table of Contents

1 Introduction

"An effective board should not necessarily be a comfortable place." This is how the UK's Financial Reporting Council (FRC) recently introduced *Guidance on Board Effectiveness 2011*. The growing interest in board performance is by no means restricted to company investors, but has become legitimized as the general public's "right to know." Specifically, there is an increased sense of entitlement regarding the accountability and transparency of quoted company boards' de facto workings. Among others, this entitlement is due to boards' underlying duties. As the "guardian of the corporate values" (Stiles and Taylor, 2001), the board plays a decisive role in a company's growth strategy. Not only does the board set the standards for the strategy process's quality and nature, but therefore also defines the boundaries of what is acceptable and what is not. Accordingly, the board is generally referred to as the top organizational decision maker as it constitutes the organization's belief system.

It is argued that an effective board can successfully fulfill its responsibilities and accomplish the role set required of it (Nicholson and Kiel, 2004). Encouraging boards to reflect on their role as a group and on their members' individual contributions is of vital importance for good corporate governance. An essential feature of such encouragement is to reconsider whether the way in which decisions are taken might ultimately affect their quality (FRC, 2011).

This chapter first seeks to illustrate the board's responsibility for the company's strategic direction, as well as the board's comprehensive catalogue of nontransferable duties with regard to Swiss Company Law (SCL) and its Code of Obligations (CO). Subsequently, effective board workings pertaining to a balanced organizational growth are illustrated. The chapter ends with concluding remarks.

2 Responsibility for the strategic direction

The board's responsibilities and underlying duties vary, in keeping with the specific country's jurisdiction. SCL is based on the parity principle.[1] The parity principle grants the board a great deal of power and, conversely, assigns significant responsibilities to each board member. Attention is primarily directed to the nontransferable and inalienable duties, as stipulated in article 716a CO, according to which the board of directors is not allowed to delegate final responsibility for certain core competences (Forstmoser, 2011), such as the ultimate management of the company. These duties, it is argued, are a pivotal aspect of the Swiss board system. Consequently, the board is a "central pillar" of corporate governance and is presumed to carry out all duties not assigned to the general shareholders' meeting (Hofstetter,

[1] The SCL forms part of the CO. The legal text is written in the three official national languages: German, French, and Italian. It must be noted that these three languages coexist equally. Bundesgesetz betreffend die Ergänzung des Schweizerischen Zivilgesetzbuches (Fünfter Teil: Obligationenrecht) as at March 30, 1911, SR 220.

2002).[2] Based on its election by the general shareholders' meeting, the board is characterized as a formal company organ. Furthermore, since corporate responsibility is purely personal and thus primarily addresses natural persons, board members might be held responsible for neglecting to properly fulfill their individual duties. Consequently, there is no responsibility release, such as a personal liability release, for board members regarding their nontransferable duties, even if board committees wholly undertake such duties.[3] The final responsibility for their nontransferable and inalienable duties therefore remains with each board member.

Accordingly, board members in Switzerland face a great challenge, especially regarding the substantial catalogue of nontransferable and inalienable duties. Each board member is in charge of diligently accomplishing a strategic, a control, and an institutional role. The following duties are listed in article 716a CO: In respect of the strategy role, the board is responsible for the ultimate management of the company, the organization's establishment, as well as the appointment and removal of the persons entrusted with management and representation. The control role involves the accounting system and the financial controls' structuring, the financial planning insofar as this relates to managing the company, and the ultimate supervision of the persons entrusted with the management – specifically regarding compliance with the law, the articles of incorporation, the regulations, and directives. Finally, the institutional role comprises preparing the business report and organizing the general shareholders' meeting, as well as implementing its resolutions, and notifying of the judge in the case of over-indebtedness.

A 2008 survey revealed that Swiss board members perceive their strategic duties to be the most important nontransferable duties.[4] This is in line with the amount of time that full board meetings spend on the three key board roles (strategy, control, and institutional) that underlie the categories of nontransferable duties: 53.0% of board meeting time is spent on its strategy role, 36.3% on its control role, and 10.7% on its institutional role. This clearly highlights that boards' primary role in quoted Swiss companies is based on managing and accomplishing strategic duties, since they devote over 50% of their time to these. In the following, the nontransferable duties relating to boards' strategy role are outlined.

---

[2] *Inalienable* serves to distinguish between the board of directors' competences and those of the general sharholders' meeting. Since the only formal higher authority qualified to revoke the board of directors' competences is the general shareholders' meeting, *inalienable* refers to an unallowable upwards delegation. Conversely, in the CO, the prevailing interpretation of *nontransferable* prohibits the board of directors from delegating decisions to the next subordinate authority, general management and, consequently, to all lower authorities (Kammerer, 1997).

[3] For additional information on board members' personal liability, see, for example, Forstmoser et al. (1996) and Von der Crone et al. (2006).

[4] The survey is based on an accessible population of 401 board members of the SMI Expanded® as at June 30, 2008. The final sample comprised 83 board members – a response rate of 20.7%. For further information on boards of directors' de facto nontransferable duties, please refer to Rüdisser (2009).

2.1 The ultimate management of the company

According to cipher 1 in article 716a CO, the board of directors has the nontransferable duty of "the ultimate management of the company and the giving of the necessary directives."[5] The board's responsibility is not only valuable in terms of balancing the business objectives and financial means, i.e. designing organizational activities on the basis of their fundamental principles, but also regarding the board's strategic directive and intervention responsibilities (Böckli, 2004). The ultimate management of the company is therefore the board of directors' core function and concerns agreement on and determination of well-defined objectives, the course of action to achieve these planned objectives, and the company's positioning in the market (Chapuis, 2002). In short, "[p]rioritizing the adoption of ever narrower financial means" (Böckli, 2004).

With regard to the implementation of strategic fundamentals, the board must provide the general management or others in charge with the necessary directives (Watter, 2002). Consequently, the board must actively guide the company's strategic direction. These directives are, however, not bound to a specific form. Various options are thus possible, such as recording board rulings, organizational regulations, and providing general management with oral or written assignments (Chapuis, 2002). The survey participants prioritized nontransferable strategic duties as follows (Rüdisser, 2009): defining strategic goals, questioning the chosen strategy, determining means to achieve the goals, and periodically checking the allocated means.

2.2 The establishment of the organization

The board is ultimately responsible for the continuing organizational set-up that meets the challenges of today's competitive environment, as well as the market conditions (Krneta, 2005). Within cipher 2 of article 716a CO, the board is given the nontransferable duty of "the establishment of the organization." The board is therefore in charge of establishing both the main organization and the organization's basic structure functionally and conveniently, for example, designating it a parent or holding company, a subsidiary or branch establishment, etc. (Kammerer, 1997). In addition, cipher 2 relates to establishing direct subordinates, i.e. the general management, but also relates to organizing the board of directors. Organizational regulations are the basis for the board's formation and are supplemented by organization charts, job descriptions, and functional specifications (Forstmoser et al., 1996).[6]

---

[5] The meaning of *ultimate management* has been ambiguous since the introduction of article 716a CO in 1992 (Nobel, 2005). Hence, to date, the legal academic community and court practice do not consistently agree on the interpretation of this law's wording. While some assign the board a comprehensive and very active role in the conduct of business, others interpret the law as a mere guideline for the board oversight role. Although *strategy* is not explicitly specified in the legal text, it is considered to be generally included in *ultimate management*.

[6] The mandatory content of the organizational regulation is subject to article 716b CO, which states that the articles of incorporation authorize the board to fully or partially delegate the management of duties other than those that are nontransferable to individual members or third parties. Regulations organize the management, determine the required positions, define further duties, and regulate reporting (article 716b paragraph 2 CO).

It is important to bear in mind that the board is allowed to delegate the elaboration of an organizational regulation (decision shaping). However, in the end, only board members decide on its content and, consequently, assume full responsibility for it (decision making) (Böckli, 2004). The survey participants prioritized the nontransferable organizational duties as follows (Rüdisser, 2009): questioning the chosen organizational form, assigning roles and responsibilities, determining internal management and internal control, as well as defining the structural and operational organization.

2.3 The appointment of the persons entrusted with management

In cipher 4 of article 716a CO, the board is assigned "the appointment and removal of the person entrusted with the management and the representation." The latter includes holders of procurement as well as other mandate holders. Again, boards are not obliged to be in charge of all human resources affairs, only those of the highest subordinate level, namely general management and directors. Secondary level appointments or dismissals can thus be delegated (Chapuis, 2002).

Owing to the multitude of authorized representatives in larger corporations, the signing of authorization rights is also subsumed under cipher 4 (Botschaft, 1983). This has led to serious discussions on whether the general management is allowed to pass on these rights after directly obtaining the signing authorization rights from the board (Chapuis, 2002). Watter (2002) argues that the latter is allowed if the board has explicitly considered each position within the organizational regulation.

In both cases, members of the board take charge of personnel policies, including personnel appointment, evaluation, reward, and development (Kammerer, 1997). The survey participants prioritized nontransferable human resources duties as follows (Rüdisser, 2009): evaluating the management, manager remuneration, finding management team members, as well as developing the management.

3 Effective workings for balanced growth

As indicated by the nontransferable duties, the board is ultimately responsible for the organization's strategic direction. However, not all boards place equal emphasis on identifying and populating strategic growth opportunities. Instead of leaving everything to management, the board could make a difference by becoming actively involved in new growth initiatives. This could sustain overall commitment and, hence, foster continuous and balanced growth.

But there is no such thing as a free lunch. In order for boards to set and frame the growth challenge, they not only need to offer strong personal commitment – such as making sufficient time available for their mandate – but should also assure effective workings while working as a group. Finally, a key factor of board processes' efficiency in terms of doing "the right things now and in the future,"[7] involves a reflective evaluation of the work performed, as well as an innovative succession planning.

In the following, these critical success factors are presented with a view to their underlying improvement opportunities.

3.1 Personal commitment as a key enabler

The increase in legislation (hard law) and governance codes (soft law) reflects board work's growing importance and complexity. Currently, it is perceived as vital that boards should demonstrate how they meet their responsibilities and carry out their roles most effectively. Each board member's personal commitment is considered a key enabler to successfully carry out board processes in the group. Given that time reflects commitment and active involvement, it could be a potential prerequisite for board work's quality. Nevertheless, there are substantial differences in the amount of time board members devote to their mandate. There are major differences between the time board chairmen and other board members devote, which could be due to their differing roles. On average, board chairmen spend more than twice the amount of time on individual preparation than non-chairman board members do (Rüdisser, 2009).

Even within companies with similar business dimensions, there are considerable differences between their regular board members' commitment, such as not being allowed to serve on more than two board committees. Board members of quoted Swiss companies, for example, indicated that they spend an average of more than 20 working days on their board mandate per year. Whereas 84.8% spend between 50 and 300 working hours on their board mandate, 11.1% work more than 300 hours per year, and 4.1% stated that they work fewer than 50 hours on their board mandate per year. The difference between the minimum value – 5 hours per board mandate per year – and the maximum value – 924 hours per mandate per year – is surprisingly wide and remains thought-provoking (Rüdisser, 2009).

Since board meetings are time-constrained and the quantity of transacted business is extensive, the ability to influence board members during these meetings is critical (Tuggle et al., 2010). To ensure that board members dedicate sufficient time to their specific duties, the *EU Corporate Governance Framework Green Paper* suggests limiting the number of board mandates as a "simple solution" (European Commission, 2011). However, it should be emphasized that this not only depends on the total amount of time that board members devote, but also depends on their individual time flexibility. Given this level of effort and substantially more personal commitment to their board activities, it can be deduced that increased preparation leads to greater knowledge, which ultimately results in improved board effectiveness.

[7] Peter F. Drucker et al.'s (2008) remarks regarding effective mission statements – such as the degree to which objectives are achieved and the extent to which targeted problems are solved – also apply to board processes' efficiency. While efficiency accounts for input-output effects in the sense of "doing the thing right," effectiveness focuses on the output – "doing the right thing."

3.2 Qualified involvement and priority setting

Another essential requirement for a board to be effective is a well-structured and elaborated decision-making process. To achieve this, essential conditions are that well-informed board members should be involved and that the prioritizing of the underlying duties should be focused. Nevertheless, some critics claim that boards tend to focus on passive, retrospective working methods instead of on more active engagement, such as vigorously initiating the search for new growth opportunities. Furthermore, it has been shown that, according to individual board members, disproportionate attention is given to financial duties in full board meetings (Rüdisser, 2009). Reasons for the discrepancy are manifold. One explanation could be that during their meetings boards spend too much time on past financial issues at the expense of other, more important information such as current market data, up-to-date risk profiles, ongoing practices and standards of internal control, as well as reports on near-term and long-term business opportunities.

In order to improve the quality of board involvement and subsequent priority setting, board members should rather meet in specific workshop settings instead of in formal, static conference rooms where they must sit through management team presentations (Rüdisser, 2010). In addition, each board member should be asked to submit additional creative ideas for the agenda and to actively contribute to the list of key strategic decisions that need to be made. In this regard, the chairman could play an important role. It is even argued that "[g]ood boards are created by good chairmen," since they could not only ameliorate the conditions in terms of the overall board effectiveness, but could also enhance board members' individual commitment by setting standards of "integrity and probity, and (…) clear expectations concerning the company's culture, values and behaviours, and the style and tone of board discussions" (FRC, 2011). Nevertheless, each board member has the opportunity to effectively influence the identification and population of the organization's growth opportunities and, to a certain extent, also bears responsibility for these opportunities.

3.3 Innovative succession planning and reflective evaluation

It is argued that the strongest and most effective boards are those with a variety of relevant perspectives (Hohn, 1997). Accordingly, board members are required to play multiple roles, "in some cases dipping deep into the details of particular business, in others playing the devil's advocate, in still others serving as the project manager" (Sonnenfeld, 2002). Reflective thinking and the continuous challenge of the underlying working procedures are very important. It is argued that the following personal attributes allow for a diversity of perspectives among board members and might ensure that the board does not consist of "like-minded individuals" (FRC, 2011): intellect, critical assessment and judgment, courage, openness, honesty and tact, the ability to listen, forge relationships, develop trust, and "the intellectual capability to suggest change to a proposed strategy, and to promulgate alternatives" (p. 10).

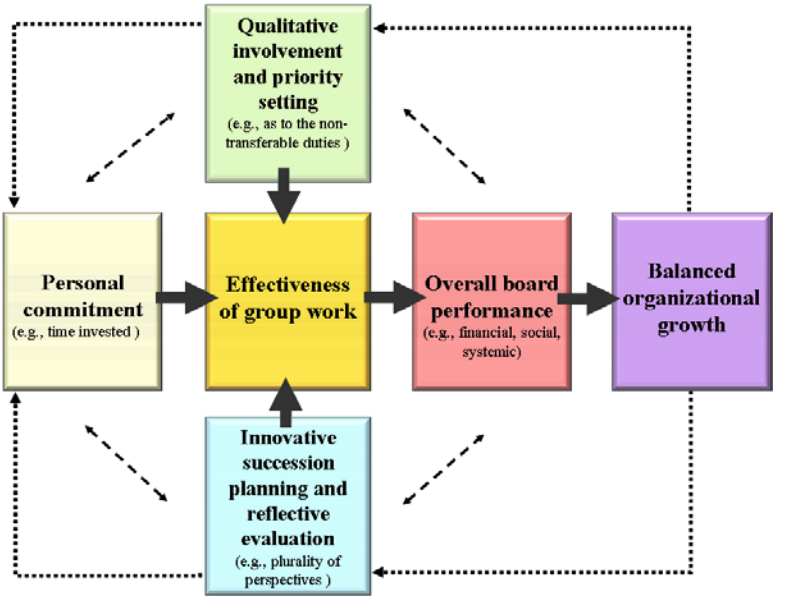
Innovative succession planning is a critical element in identifying those board members who are willing and able to make a positive contribution, and, consequently, enhance board effectiveness. Nevertheless, people often tend to favor behavior patterns similar to their own, which results in fairly homogeneous boards that impede a variety of thoughts. Instead of board members being comfortable with well-tried and traditional search methods, they could "think out of the network," which might be a favorable and worthwhile option for the board's succession planning. In fact, heterogeneity is considered a future competitive advantage for a company.

Board evaluation has become an integral part of good corporate governance, given that a rigorous performance appraisal is the basis for effective board work. Board members are therefore urged to reflect on how to continually improve their performance by maximizing their strengths and highlighting areas for further development. This can be achieved by introducing appropriate development initiatives. Currently, many countries have introduced of regulations that require regular board performance reviews.

The European Commission, for instance, recommends that it should be a role of non-executive or supervisory directors of listed companies to evaluate the board's performance annually, which should include assessing its membership, its organization and operation as a group, the competence and effectiveness of each board member and board committee, and how well the board has performed against a set of performance objectives (European Commission, 2005). In addition, this Green Paper advises boards to make use of an external facilitator to add value to board evaluation by sharing best practice experiences and allowing an objective perspective. The review should encourage openness, as well as maintain a degree of confidentiality, while covering the quality and timeliness of information received by the board, the management's response to requests for clarification, and assessing the chairman's role (European Commission, 2011).

3.4 Summary

The following figure illustrates the abovementioned relationships.



(Source: Author)

Figure 1: Boards' contribution to balanced organizational growth

4 Conclusion

It is time to overcome the general fear of non-compliance, which often gives rise to an exaggerated focus on the board's control role. Boards need to – once again – be management's true strategic partner. With the aim of regaining balance in the board activities, the directors should be encouraged to reflect on their role as a group, as well as on their individual contributions to their mandate. Accordingly, contemporary board members should be actively involved in the company's strategy – not only to fulfill their legal obligation but, more importantly, to make a significant contribution to the company's future.

In order to foster organizational growth, it is crucial that board members demonstrate strong personal commitment. This implies that sufficient time will be made available for each individual board mandate and, while working as a group, effective working procedures are assured. To improve the quality of board involvement and their subsequent priority setting, each board member should be asked to regularly submit innovative ideas for the board's agenda and should contribute to the list of key strategic decisions that need to be made. The emphasis should not be on mere augmentation of growth percentages, but rather on having an enduring and sustainable impact.

A further critical success factor to enhance board effectiveness is innovative succession planning. Heterogeneity is specifically considered a future competitive company advantage. A favorable and worthwhile option for the board's succession planning is therefore to "think out of the network" instead of being comfortable with well-tried and traditional search methods. Furthermore, rigorous performance appraisal is crucial for effective board work. Consequently, board evaluation has become an integral part of good corporate governance. By maximizing strengths and highlighting areas for further development, board members are continually urged to reflect on how to improve their working procedures as well as their underlying performance.

This shift might also bring a welcome debate on how boards can effectively contribute to balanced organizational growth. Arguments about which board model or structure is superior seem to be outdated and overshadowed by more relevant questions, such as: How can a board member make the board work best for the company and its stakeholders? Boards have the potential to directly influence companies' growth strategies. It is a matter of tapping their full potential. For better or worse, the "tone at the top" continues to be a crucial factor in finding strategies for sustainable development of balanced organizational growth.

References

Böckli P. (2004). *Schweizer Aktienrecht. Mit Fusionsgesetz, Internationalen Rechnungslegungsgrundsätzen IFRS, Börsengesellschaftsrecht, Konzernrecht und Corporate Governance*. Zürich: Schulthess.

Botschaft. 1983. Botschaft des Bundesrates über die Revision des Aktienrechts vom 23. Februar 1983. *Bundesblatt* 1983 II.

Chapuis C. (2002). Art. 716a. In J. Kren Kostkiewicz, U. Bertschinger, I. Schwander, P. Breitschmid (Eds.), *OR Handkommentar zum Schweizerischen Obligationenrecht* (pp. 870–875). Zürich: Orell Füssli.

Drucker P. F. (1993). *The Five Most Important Questions You Will Ever Ask About Your Nonprofit Organization: Participant's Workbook*. San Francisco: Jossey-Bass.

European Commission (2011). Green Paper on the EU Corporate Governance Framework as of April 5, 2011. Brussels: European Commission. http://ec.europa.eu/internal_market/company/docs/modern/com2011-164_en.pdf.

European Commission (2005). Commission Recommendation 2005/162/EC of 15 February 2005 on the role of non-executive or supervisory directors of listed companies and on the committees of the (supervisory) board. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:052:0051:0063:EN:PD.

Forstmoser P. (2011). *Organisation und Organisationsreglement der Aktiengesellschaft. Rechtliche Ordnung und Umsetzung in der Praxis*. Zürich: Schulthess.

Forstmoser P., Meier-Hayoz A., Nobel P. (1996). *Schweizerisches Aktienrecht*. Bern: Stämpfli.

FRC (2011). Guidance on Board Effectiveness 2011. London: Financial Reporting Council. http://www.frc.org.uk/images/uploaded/documents/Guidance%20on%20board%20effectiveness%20FINAL5.pdf.

Hofstetter K. (2002). *Corporate Governance in Switzerland. Final Report of the Panel of Experts on Corporate Governance*. Zürich: Economiesuisse.

Hohn H. G. (1997). Board Governance: Balance is the Key. *NACD Directorship*, 23(4), 1-3.

Kammerer A. (1997). *Die unübertragbaren und unentziehbaren Kompetenzen des Verwaltungsrates*. Zürich: Dissertation, Rechtswissenschaftliche Fakultät der Universität Zürich.

Krneta G. (2005). *Praxiskommentar Verwaltungsrat. Art. 707–726, 754 OR und Spezialgesetze. Ein Handbuch für Verwaltungsräte*. Bern: Stämpfli Verlag.

Nicholson, G. J., Kiel, G. (2004). A Framework for Diagnosing Board Effectiveness. *Corporate Governance*, 12(4), 442–460.

Nobel P. (2005). Podiumsdiskussion. Ist Art. 716a OR "lettre morte?" In C. M. Baer (Ed.). *Verwaltungsrat und Geschäftsleitung. Ihre Tätigkeit und ihr Verhältnis zueinander. Tagungsband zum 10. Aktienrechts-Forum 2005* (pp. 101–108). St. Gallen: Haupt.

Rüdisser M. F. (2010). Rethinking Board Communication: How to Enhance Critical Reflection within the Boardroom. In Haunreiter, D. (Ed.), *Kommunikation in Wirtschaft, Recht und Gesellschaft* (pp. 3-23). Bern: Stämpfli.

Rüdisser M. F. (2009). *Boards of Directors at Work: An Integral Analysis of Nontransferable Duties under Swiss Company Law from an Economic Perspective*. Bamberg: Dissertation 3657, University of St. Gallen.

Sonnenfeld J. A. (2002). What Makes Great Boards Great. It's not Rules and Regulations. It's the Way People Work Together. *Harvard Business Review*, 80(9), 106-113.

Stiles P., Taylor B. (2001). *Boards at Work. How Directors View their Roles and Responsibilities*. New York: Oxford University Press.

Tuggle C. S., Schnatterly K. Johnson R. A. (2010). Attention patterns in the boardroom: How board composition and processes affect discussion of entrepreneurial issues. *Academy of Management Journal*, 53(3), 550–571.

Von der Crone H. K., Carbonara A., Hunziker S. (2006). *Aktienrechtliche Verantwortlichkeit und Geschäftsführung. Ein funktionaler und systematischer Überblick*. Basel: Helbling & Lichtenhahn.

Watter R. (2002). 26. Titel: Die Aktiengesellschaft. In H. Honsell, N. P. Vogt, R. Watter (Eds.). *Basler Kommentar zum Schweizerischen Privatrecht, Obligationenrecht II* (pp. 915–957). Basel: Helbling & Lichtenhahn.

Part C

Balanced Growth – a resource management perspective

HR strategies for balanced growth

David J. G. Dwertmann and Justus Julius Kunz

Abstract:

In order to fulfill the expectations of shareholders, growth has become an inevitable imperative for most companies. However, growth is always associated with increased complexity. Additional resources, especially human capital resources, are necessary to master this added complexity. Companies can satisfy this demand by either recruiting new employees, or by increasing their current personnel's efficiency. However, in times of extreme talent bench shortfall and rapidly increasing burnout statistics, HR departments are facing serious challenges. Firms may not be able to recruit the necessary talent, may dilute the company's culture and identity through intense recruitment within short periods, or may overload their current personnel quantitatively or qualitatively. This chapter outlines how companies can master these challenges through strategies such as: expanding the recruitment population, becoming a desired employer, recruiting very deliberately, prioritizing tasks and services, and, finally, interlinking strategic human resource management and high performance work systems' activities. By applying these strategies, companies can avoid the risks of excessive growth. Instead, they can capitalize on times of growth, establishing sustainable or balanced growth in order to get ahead of their competitors.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_10,

Table of Contents

1 The importance of HR

"No country, no organization can remain competitive unless talent – the engine force of economies – is there to ensure success of organizations in turbulent times [...]."[1] Considering this statement, which reflects many scholars and practitioners' beliefs, it seems alarming that human capital resources specifically are becoming increasingly scarce (McKinsey, 2011; Olson et al., 2008). The workforce is becoming a more important source of competitive advantage (Kroth and Keeler, 2009; Pfeffer, 1994) than traditional resources (i.e. financial resources, technology, and economies of scale). Consequently, so is long-lived or balanced firm growth, which seems to be an inevitable imperative nowadays (Christensen et al., 2003). Thus, it can be said that "human capital is replacing financial capital as the engine of economic prosperity" (World Economic Forum, 2011, p. 7).

In line with these statements, the surveyed Chief Executive Officers in PricewaterhouseCooper's 14th Annual Global Survey (PWC, 2011) listed the "the race for talent" and "a lack of key skills" as their top priorities. According to Michaels and colleagues (2001), this "war for talent" started in the late 1990s, prior to the bursting of the dot-com bubble. However, the demand for human capital resources, including "the training, experience, judgment, intelligence, relationships, and insight of individual managers and workers in a firm" (Barney, 1991, p. 101), was only temporarily reduced by the following crises (the dot-com crisis, financial crisis) and remains a concern for companies to this day (PWC, 2011).

This increasing importance of human capital resources for firms and their growth is grounded in at least five different sources: the irreversible shift from the industrial age to the information age (Michaels et al., 2001); the intensifying demand for high-caliber managerial talent due to the new job complexity arising from globalization and rapid advances in technology (Michaels et al., 2001); insufficient educational standards in most countries (World Economic Forum , 2011); employees' growing propensity to switch companies (Michaels et al., 2001); and the demographic change in many industrialized countries (World Economic Forum, 2011).

Owing to the increasing importance of human capital resources for company growth, we will review the interplay of organizational growth and human capital resources in this chapter. The following questions will be raised and answered: What challenges do companies, and especially Human Resource Management (HRM) departments, face in times of growth? What solutions do HRM departments offer for the challenges in times of growth? Which risks are associated with each of the alternatives? And, finally, how can the HRM department reduce these risks to help the company gain a sustainable competitive advantage (SCA) and establish a sustainable or balanced growth?

[1] Jean Charest, Premier of Québec, Canada in the foreword of the World Economic Forum (2011).

2 HRM in times of growth

Times of growth are usually associated with change and are, therefore, challenging for organizations. There are two options regarding how to grow (Hagel, 2002): companies can grow by acquiring other firms (e.g., Christensen et al., 2011), or they can rely on organic expansion (e.g., Hagel, 2002). Organic expansion means that firms grow internally by acquiring new and more projects (or developing new products), or by extending the volume of their current projects (or the sales of their current products). Both options, acquisitions and organic expansion, lead to more complexity (Penrose, 1959) and, thus, require additional resources (Zott, 2003). Whereas in acquisition scenarios, organizations often simultaneously acquire new complexity and new personnel to support the existing staff (from the former firm), organic growth as such comes without new external personnel. Thus, organizations have to react to these new challenges.

From an HR perspective, there are two alternatives to react to the increased requirements and demand for qualified personnel in times of organic growth: companies can either hire new employees to handle the additional tasks, or they can maintain their current staff and try to make them work more efficiently (see Figure 1). However, each of the alternatives is associated with two main risks and, therefore, challenges. Only if companies, and especially HR departments, are capable of dealing with these challenges, can they continue to grow, gain an SCA and, thus, establish balanced growth.

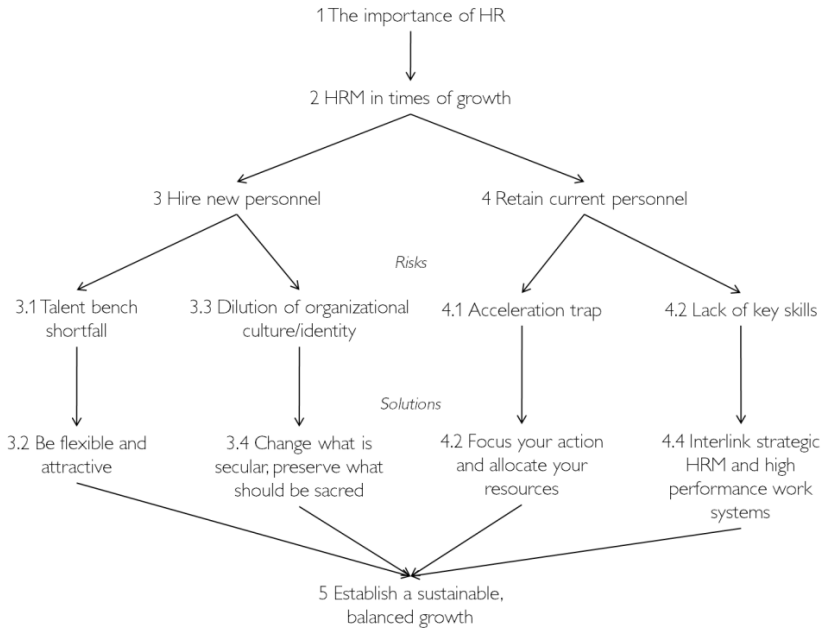


Figure 1: Overview of the chapter: HR strategies for balanced growth

3 Hire new personnel

The recruitment of new employees is the first option for HR departments in times of organic growth. This option is indicated if the current staff is already working to full capacity. However, two main risks are associated with reliance on the recruitment of new personnel, which are therefore challenges for the HR department. First, there is no longer a guarantee that qualified candidates will be found and attracted. Second, while extensive hiring of new employees can refresh an organization's innovativeness and culture, it can also threaten the positive aspects of the existing corporate culture and identity, which can be a source of SCA (Barney, 1986).

3.1 Talent bench shortfall

Companies can rely on the external job market to recruit the personnel they need for organic growth. However, this approach seems risky since firms might no longer find appropriate employees. This failure is due to, on the one hand, jobs becoming increasingly complex (van Yperen and Hagedoorn, 2003) and requiring highly skilled staff-members. On the other hand, the talent pool of young and well-educated workers is shrinking in many countries (Michaels et al., 2001). In the words of Jean Charest, Premier of Québec, Canada: "We are entering the era of unparalleled talent scarcity, which, if left unaddressed, will put a brake on economic growth around the world, and will fundamentally change the way we approach the workforce challenges."[2]

Both the mentioned effects – increasing job complexity and a shrinking pool of talent – are especially relevant regarding western industrialized countries. In times of globalization, many jobs in the manufacturing and production area are outsourced to low-earning countries (Farrell and Grant, 2005). Therefore, the remaining jobs are more demanding and complex than before. This development calls for better education for the youth and more training for young employees, which often do not happen (World Economic Forum, 2011). However, not only western countries are facing this problem. China, for example, also needs to pay attention to the education of its youth if it wants to go on growing. There is currently already a shortage of young and well-educated employees in China who can meet multinational companies' job demands (Beechler and Woodward, 2009).

The second issue – the shrinking pool of young and talented people – is also very threatening in respect of western industrialized countries (Boehm et al., 2011). This issue is caused by the major demographic change trend. Three sources are relevant with regard to the demographic change: first, life expectancy has increased; second, fertility rates have declined in many countries; third, rates of migration have influenced countries' populations. Germany is an extreme example of a country affected by demographic changes, therefore its statistics will be used to outline the interrelationship of the three sources.

[2] In the foreword of the World Economic Forum (2011).

The German Statistics Office (Deutsches Statistisches Bundesamt, 2009) forecasts two scenarios for the development of life expectancy: until 2060, the life expectancy of men will either increase by 7.8 years, and that of women by 6.8 years, or, if the health sector makes extensive progress, the increase will be 10.6 years for men and 8.8 years for women. The second scenario would result in an average life expectancy of 87.7 years for men and 91.2 years for women.

In many countries, fertility rates dropped dramatically in the second half of the 20th century. In Germany, the fertility rate was only 1.37 in 2007 (Schweizer Bundesamt für Statistik, 2009). It is important to note that countries would need a fertility rate of 2.1 children per woman just to maintain the population level.

Finally, the migration rate influences demographic change. This source of change is currently supported by national governments removing legal and regulatory hurdles to increase international interaction (Beechler and Woodward, 2009). A positive migration rate value indicates that more people are moving to the country than are leaving. In this respect, Germany has experienced more immigration than emigration since 1985 (Deutsches Statistisches Bundesamt, 2010a). However, the positive value of the migration rate has decreased steadily since 1992[3]. In sum, migration cannot compensate for the increasing life expectancy and the low fertility rate's negative effects. This leads to a demographic change; that is, population aging and population decline – specifically a shrinking workforce

The described effects and additional sources (e.g., the retirement of the so-called baby boomer generation, see, Wey Smola and Sutton, 2002) lead to a shrinking workforce and a resultant shift of power from the company to the individual (Michaels et al., 2001). Today, well-educated university graduates can often choose between different appealing job opportunities (Ray and Jeon, 2008; Van Hoye and Saks, 2011). This shift in power has also led to higher fluctuation rates of employees between companies (Ng et al., 2007).

All these developments have led to companies finding it harder to recruit new employees. Olson et al. (2008) have identified the talent bench shortfall as one out of four major reasons for revenue stalls (negative growth)[4]. In Germany, nine out of ten companies are expecting a talent bench shortfall in the next decade (McKinsey, 2011). However, if companies could recruit the necessary employees, they could gain a competitive advantage and, thus, support their growth. There are multiple ways in which organizations can cope with this challenge and turn the potential threat into a strength.

3.2 Be flexible and attractive

One way to avoid the competition for young and well-educated talents is to simply change the target group. Companies could consider hiring older employees (McKinsey, 2011). In Germany, less than 50% of companies employed people

[3] Own calculations based on the Deutsches Statistisches Bundesamt (2010a) and the Deutsches Statistisches Bundesamt (2010b).

[4] See also McKinsey (2011).

older than 50 in 2001 (Bellmann and Kistler, 2003). Indeed, older employees are often stigmatized at the workplace (Loretto and White, 2006; Van Dalen et al., 2010) due to general expectations regarding declining performance over the lifespan and fluid intelligence's research results (e.g., Salthouse 2009). However, contrary to public opinion and intelligence research, scholars have shown evidence that "refutes the notion of a universal negative relationship between age and job performance"[5] (Kanfer and Ackerman, 2004, p. 440). Recent studies by Malmberg et al. (2008) and van Ours and Stoeldraijer (2010) even suggest that particular age groups' contribution to company productivity increases with age until they are 40-45 years old and thereafter remains stable. In their meta-analysis, Ng and Feldman (2008) suggested that the conflict between general opinion and research findings could have been caused by solely focusing on core tasks, instead of taking outcomes on the environment into account (e.g., general counterproductive work behavior). In sum, older employees can certainly be a valuable resource for firms searching for new employees.

Another way of coping with the changed situation on the labor market is to depict the company as an attractive employer (World Economic Forum, 2011). Participation in various initiatives like Top Job in Germany (www.topjob.de) or the Great Place To Work initiative in Switzerland (www.greatplacetowork.ch) and other countries is an example of an attempt to position the own company ahead of other competitors for talent. This relatively new approach (Backhaus and Tikoo, 2004) is called employer branding (or employer attractiveness).

According to Backhaus and Tikoo (2004, p. 501) "Employer branding represents a firm's efforts to promote, both within and outside the firm, a clear view of what makes it different and desirable as an employer." In a similar vein, Berthon and colleagues (2005, p. 151) state that "Employer attractiveness is defined as the envisioned benefits that a potential employee sees in working for a specific organisation. It constitutes an important concept in knowledge intensive contexts where attracting employees with superior skills and knowledge comprises a primary source of competitive advantage."

Today, almost every large company engages in employer branding. However, the question is: How can companies become known as an attractive employer brand? One of the important factors is that firms need to differentiate themselves from their competitors (Backhaus and Tikoo, 2004). In an industry where all firms are regarded as reliable, a company may want to be known as the most innovative one. However, if every competitor in the industry is perceived as extremely innovative, a firm may want to convey an image as a very reliable employer. A second important factor is the simultaneous application of internal and external employer branding strategies (Backhaus and Tikoo, 2004). Internal strategies, for example, include compensation, the work environment, and the leadership culture, whereas external strategies include image campaigns, sponsoring, and employment ads (e.g., Schumacher and Geschwill, 2009).

[5] For example, see McEvoy and Cascio (1989) and Waldman and Avolio (1986).

According to Backhaus and Tikoo (2004), in the HR practitioner literature, employer branding is described as a three-step process. First, develop a value proposition based on the organization's culture or identity, the management style, current employees, and your product. This value proposition should be a true representation of the value that the firm offers its employees. In this context, the firm shows how it differs from its competitors. One way to investigate this value could be employee attitude surveys or customer surveys. Second, the firm needs to consistently communicate the value proposition to potential recruits, placement counselors, etc. The external marketing should attract the target population and enhance the product and/or the corporate brand. Finally, the internal marketing should guide recruits and other employees to internalize the company value proposition in order to retain all employees (Frook, 2001).

The recruitment of new employees is essential for companies in general (Boxall, 1996) and especially in times of growth. In the words of Olson et al. (2008, p. 58), the talent bench shortfall "has become a fact of daily life in many industries and functions. Indeed, at this writing, shortages of critical talent are the primary concern of human resources departments globally, not just in high-growth markets but in a range of specialty skill categories, and they are expected to get worse. What stops growth dead in its tracks, however, is not merely a shortage of talent but the absence of required capabilities – such as solutions-selling skills or consumer-marketing expertise – in key areas of a company, most visibly at the executive level." The short answer to this challenge is strategic HRM[6]. However, even if companies engage in strategic HRM and do not have problems recruiting the necessary talent in times of growth, another risk threatens managers who want to achieve balanced growth – the dilution of the positive aspects of the organizational culture/identity.

3.3 Dilution of organizational culture/identity

Even if companies can recruit the necessary personnel in times of growth, they are not immune to problems associated with human capital resources. In times when growth is not balanced, but instead excessive, the need for new resources might lead to the hiring of a large number of new employees in a short time. This is not only an opportunity to refresh old thinking, but also carries the risk of diluting the company's culture and the identity. The people make the place (Schneider, 1987), and thus, in times of intense recruitment, a place can change in an undesired direction (DeLong et al., 2008). However, companies that can preserve the positive aspects of their culture and identity, can expect various desirable outcomes, such as increased commitment and decreased turnover (Cole and Bruch, 2006).

Barney (1986) produced a broad definition of organizational culture that suits the application here. He states that "[In this work,] organizational culture [is] typically defined as a complex set of values, beliefs, assumptions, and symbols that define the way in which a firm conducts its business. In this sense, culture has pervasive effects on a firm because a firm's culture not only defines who its relevant

[6] See, for example, Schuler and Jackson (2007).

employees, customers, suppliers, and competitors are, but it also defines how a firm will interact with these key actors" (Barney, 1986, p. 657).

According to Gioia and Thomas (1996), "corporate or organizational identity concerns those features of the organization that members perceive as ostensibly central, enduring, and distinctive in character that contribute to how they define the organization and their identification with it" (p. 372).

Thus, both concepts refer to central organization characteristics that, on the one hand, guide employees on how to behave and, on the other hand, determine how employees relate to the company. Organizational culture and organizational identity develop over time and can be a source of SCA (Barney, 1986). In this respect, the terms culture and identity are used similarly, because the arguments made refer to both concepts.

Barney (1986) identifies three different characteristics that need to be fulfilled for organizational culture to provide an SCA: culture needs to be valuable, rare, and imperfectly imitable. The condition of being valuable is met if the culture enables the firm to generate lower costs or higher income. For example, if there is a very positive climate in the company, each employee feels valued, and is given the opportunity to develop, the firm may have a valuable culture, because the turnover rate is very low (e.g., Sheridan, 1992). This leads to less spending on recruitment[7] (see employer branding in Section 3.2). Rareness means that the organizational culture is not common to every competitor in the industry. A positive culture requires effort from the top management and develops over time. In many industries, a culture of appreciating each other is therefore very rare. Finally, the culture needs to be imperfectly imitable to lead to an SCA. How do you describe an organizational culture and, even more challenging, how do you determine what aspects of a particular culture adds value to the firm? These questions may be very hard to answer[8] (Lippman and Rumelt, 1982). Of course, if you cannot identify those critical aspects, it is not possible to imitate a firm's culture. Therefore, organizational cultures may often be imperfectly imitable.

If, through transformational leadership (Bass and Avolio, 1993), teaching processes (Schein, 1983), etc.[9], companies have managed to achieve a certain positive culture or strong identity that acts as a SCA, the question is: How stable is this culture? According to Schein (2010, p. 219) there are three main sources of organizational culture: "(1) the beliefs, values, and assumptions of founders of organizations; (2) the learning experiences of group members as their organization evolves; and (3) new beliefs, values, and assumptions brought in by new members and new leaders." Thus, in times of extensive growth, when large numbers of new employees enter the organization, the third source may change the organizational culture/identity and influence it in an undesired direction (DeLong et al., 2008).

---

[7] In general, turnover seems to be negatively associated with business unit and organizational success (Glebbeek and Bax, 2004; Kacmar et al., 2006; McElroy et al., 2001; Shaw et al., 2005).

[8] For a detailed elaboration, see Barney (1986).

[9] For a detailed analysis, see Schein (2010).

3.4 Change what is secular, preserve what should be sacred

How can organizations foster the adaptation of the organizational culture/identity and reduce the contagion by new, undesired influences? There are multiple ways: companies can include personal values and beliefs into their set of recruitment criteria. Instead of solely focusing on skills, education, and performance, soft factors can make a difference in how well an employee fits into a firm. Schneider's (1983) attraction-selection-attrition model even proposes that organizations automatically evolve toward interpersonal homogeneity. People are attracted to organizations that are similar to themselves. Organizations hire people who fit into them. When people enter the organization, the process continues and people who do not fit in, leave the organization. However, if the pressure to find new employees is too strong, companies may hire candidates who would not have been accepted under conditions of deliberate and balanced growth. Those recruits can either influence the organizational culture negatively, or they just do not fit in, will always be a foreign body, underperform and leave the company soon. All these outcomes are undesired and expensive for the organization. Therefore, companies should keep these effects in mind and preserve their core values and identity by recruiting responsibly.

By applying such sustainable and balanced procedure, companies can support their successful long-term growth. It is noteworthy that recruiting people who fit in with the company's culture does not refer to employee diversity or change in general[10]. However, "[t]ruly great companies understand the difference between what should never change and what should be open for change, between what is genuinely sacred and what is not" (Collins and Porras, 1996, p. 66).

This deliberate recruitment is an important starting point to prevent the dilution of the positive aspects of the organizational culture/identity, to instead capitalize on the new ideas that the new personnel contribute. Nevertheless, in general, firms need to monitor the development of their recruits and employees over time. On the one hand, recruiting decisions can simply go wrong. People can engage in impression management (Barrick et al., 2009; Bolino et al., 2008), especially if the organization values are actively communicated outside the company. In addition, what people say does not necessarily reflect their actions. Recruits may have matching values, but they may also be unable to apply them. On the other hand, employees' values and beliefs can change over time. Certain events at their work and also in their private lives can influence people strongly. Monitoring can become a routine by integrating behavior reflecting the firm's core values and culture into performance-based benefits and discussing this in a regular appraisal interview. It is important to be strict about these soft facts. A specific manager or employee should not be allowed to violate the company rules just because this person outperforms his or her colleagues. Only if all the employees believe that the management is actually serious about the organizational culture and values, will they act accordingly. "Values initiatives have nothing to do with building consensus – they're about imposing a

[10] Which was shown to be beneficial under certain conditions like task complexity (e.g., Bowers et al., 2000; Wegge et al., 2008).

set of fundamentally, strategically sound beliefs on a broad group of people" (Lencioni, 2002, p. 117).

The beliefs, values, and assumptions of organization founders have the strongest influence on the organizational culture at the time of founding (Schein, 2010). However, even in later stages of development, key persons in the organization still have a massive impact. Leaders act as role models[11] for their direct followers, as well as for other employees (Brown and Treviño, 2006; Davis and Rothstein, 2006)[12]. Therefore, the top managers of the company have to lead by example and behave as they expect their employees to behave. The important role of leaders is also emphasized by Badaracco (1998, p. 120), who states that "[m]anagers need to determine if their ethical vision will be supported by their coworkers and employees."

Finally, mentoring can be used to pass the organizational culture on to newly recruited employees (Chao et al., 1992; Wilson and Elman, 1990). Defining mentoring is complicated. This complexity is due to the mentoring construct being used in various fields of research (e.g., psychology, education, and management). Jacobi (1991) provided 15 different definitions of mentoring. However, according to Allen and Eby (2010), there are four common attributes in every definition: "First, mentoring reflects a unique relationship between individuals. (...) Second, mentoring is a learning partnership. (...) Third, mentoring is a process, defined by the types of support provided by the mentor to the protégé. (...) Fourth, a mentoring relationship is reciprocal, yet asymmetrical" (Allen and Eby, 2010, p. 10). Zachary (2005) finds that it is crucial to embed mentoring into the organizational culture. She states that mentoring has several positive outcomes, such as establishing ownership, promoting a shared responsibility, maintaining integrity, etc. Thus, mentoring can foster the organizational culture and should therefore be established in the company's culture.

All four outlined activities – deliberate recruitment, regularly monitoring, leading by example, and mentoring – can help companies preserve their culture in times of extensive recruitment. This is important, because "[c]ompanies that enjoy enduring success have core values and a core purpose that remain fixed while their business strategies and practices endlessly adapt to a changing world" (Collins and Porras, 1996, p. 65). However, these activities do not only have positive effects for new recruits. For example, mentoring in general promotes individual and organizational growth (Zachary, 2005). It can be used as a successful tool to reduce turnover rates (DeLong et al., 2008). Thus, the mentioned strategies benefit new employees, the whole company, as well as tenured members and can support balanced growth.

4 Maintain the current personnel

Assuming that the "war for talent" will become or already constitutes a significant challenge for companies, it is mandatory to analyze how corporate growth can also

[11] For a general description of these learning processes, see Bandura (1977; 1986).

[12] Another name for the role-modeling process is the cascading effect (Bass, 1990; Bass et al., 1987).

be managed with the existing personnel. Even if companies do not fear the dilution of their corporate culture/identity and follow the recommendations presented in the previous sections to combat the talent bench shortfall, there might be situations when suitable talent is hard to come by. Times of growth will then have to be handled with the current staff. Additionally, corporate growth is not fully predictable; consequently, hiring new people too early might lead to excessive costs and is a risky strategy – especially for smaller companies. This risk can be avoided if companies succeed in managing initial growth phases without relying on new personnel.

While maintaining the current personnel in times of growth might be the only option for some companies – either because new talents cannot be found, or because they fear the risk of accumulating too much human resource slack[13] –, this likewise poses two major risks for companies: overwhelming the employees with quantitative demands without providing the necessary resource and/or being unable to exploit growth opportunities due to insufficient training of the existing personnel. The second risk can be labeled qualitative overloading.

An excess of quantitative demands is likely to lead companies into the "acceleration trap." On the other hand, qualitative overloading results from a lack of training, specifically not monitoring whether employees possess the skills needed to handle the increase in complexity that comes with organizational growth and exploiting growth opportunities (e.g., entering new markets). Both risks endanger companies' competitiveness and overall performance. Not only might growth opportunities pass unexploited, but companies might even lose current staff members in the process if they feel overwhelmed by the increase in quantitative and/or qualitative demands. Excessive demands can result in emotional exhaustion, which might lead to burnout or a drop in retention rates.

That said, it is important to stress that both risks can be minimized if they are managed actively. Therefore, in the following, the two major pitfalls of growing with the current personnel will be explained in more detail and strategies will be introduced to minimize the risks for the company.

4.1 Acceleration trap

In times of corporate growth, companies are most likely to increase the number and speed of their activities and projects. The volume of the existing day-to-day business is either expanding, or new business areas are being exploited. If this progression is not supported by the hiring of additional personnel or the freeing up of resources, and the company does not have excess human resource slack at its disposal, the majority of the staff members will at a certain point be driven to the edge of their capacities.

Such periods of extensive growth might temporarily lead to exceptional results. In the long run, however, they do not only endanger sustainability, but also jeopard-

[13] "Slack" is a term coined by March and Simon (1958) to describe excessive resources within an organization. Bourgeois (1981) defines slack as the delta of available resources within an organization and the required resources for the normal course of business.

ize the company's competiveness (Zaugg and Thom, 2003). Companies that constantly try to increase the number and/or speed of projects beyond the point where they exceed their available resources are caught up in an "acceleration trap" and face dramatic consequences. Alarmingly, empirical research on the topic demonstrates that in the current times of globalization, cost-cutting, and shareholder-orientation, a significant number of companies are already trapped and are suffering the consequences. Therefore, it is crucial that an executive can detect over-acceleration at an early stage and can initiate effective measures to achieve balanced growth instead of organizational burnout.

A certain common pathology characterizes the majority of companies that experience chronic companywide over-acceleration and differentiates them from those that show sustainable growth. In most cases, opportunities that are initially regarded as major growth opportunities are the starting point of a negative spiral of events. Eager to capitalize on an identified chance and to boost the company performance, executives demand too many simultaneous activities with a lack of sufficient resources. While research shows that this might work well during what Bruch and Vogel (2011) call "a stretch phase of intensive energy" (p. 142), executives often harbor a serious misconception of such a phase. Believing that working beyond the company's resource limits can become the new normal, they start to constantly act on the edge of their staff's capabilities. Consequently, the overworked staff members are increasingly unable to cope with the management's increasing demands and fail to fully fulfill them. Bruch and Menges (2010) find that executives tend to misinterpret these first signs of over-acceleration as "laziness or unjustified protest" (p. 82) and react by increasing the pressure even more; thereby, creating a vicious circle that leads the company deeper and deeper into the acceleration trap.

According to Bruch and Menges (2010) and Bruch and Vogel (2011), the acceleration trap can manifest itself in three different forms that can also occur simultaneously: overloading, multiloading, and perpetual loading.

Overloading is defined as a state in which the sheer volume of activities employees have to handle exceeds the available resources. Multiloading, on the other hand, is defined as a state in which employees are confronted with multiple tasks whose priorities are not clear to them, or which are perceived as being misaligned, leaving them and their activities unfocused. Perpetual loading is a state of constant change and/or overwhelming demands without any hope of a consolidation phase. Companies caught in this form of over-acceleration are the hardest on their staff. As Bruch and Vogel (2011) emphasize, "just about anyone can tolerate overloading or multiloading for a while, especially if there's an end in sight, but when leaders neglect to call a halt to periods of furious activity, employees feel imprisoned by the debilitating frenzy" (p. 147).

The negative consequences for companies caught in one or more forms of the acceleration trap are severe. Using empirical data from 104 German companies, Bruch and Vogel (2011) show that from the employees' point of view, the turnover intention triples, while employees experience twice the aggressive and destructive behavior than that reported by members of companies not caught in the acceleration trap. This is accompanied by significantly higher levels of emotional exhaustion

and resignation. The negative consequences for the employees will often manifest themselves in increased absenteeism, higher error rates, lower adherence to delivery dates, lower levels of innovation, higher project outsourcing costs, and, ultimately, decreased customer satisfaction. The data provided by Bruch and Vogel (2011) on the acceleration trap's consequences from the top management's point of view seem dramatic but logical. Over-accelerated companies have significantly lower employee productivity, lack efficiency, have reduced growth rates, and their overall financial performance is on average a quarter below that of companies not caught in the acceleration trap. These figures demonstrate the double negative consequences of over-acceleration. It not only weakens the company in the short term by preventing it from benefitting from growth opportunities, but the effects of retention problems and disappointed customers, which ultimately "threaten the brand," are even more harmful (Bruch and Menges, 2010, p. 82).

Given the negative consequences of the acceleration trap, the relevance of the phenomenon should not be understated. Bruch and Menges (2010) specifically studied a large sample of companies (n = 600) to understand the acceleration trap phenomenon. 35% of the companies that were trapped suffered from overloading, 35% from multiloading, and 30% were in a state of perpetual loading. In total, every second company was caught in the acceleration trap. These figures provide a possible explanation for a worrying trend that the German health insurance provider BKK highlighted in a recent study on the reasons of absenteeism. Between 2004 and 2009, the number of work days lost per 1000 insured members due to the burnout syndrome had increased more than tenfold (BKK, 2011). The Hamburgische Weltwirtschafts-Institut (HWWI) even calculated that the German economy is losing 364 billion Euros due to employees suffering burnout symptoms (Meinert, 2011).

4.2 Focus your action and allocate your resources

What lesson does the research on the acceleration trap teach? Do times of growth automatically call for the recruitment of new personnel, since becoming trapped is inevitable? It is difficult to derive universally applicable recommendations. If an automotive manufacturer with a streamlined lean production system is confronted with a sudden increase in market demand for a certain model and has to double its production capacities, hiring new line worker might be the only options to avoid falling victim to the acceleration trap. However, in many other cases, the research by Bruch and Menges (2010), as well as Bruch and Vogel (2011) identify measures to combat and/or avoid over-acceleration without solely relying on hiring new personnel as the only option. The key strategy here is active resource allocation combined with a culture of focused action.

As a key success factor to master sustainable growth without solely relying on hiring new personnel in the context of escaping and avoiding the acceleration trap, active resource allocation is specifically characterized by an effective and efficient project management. In a nutshell, this means identifying the key projects that drive the company's success and ensuring that they are carried out as efficiently as possible while eliminating those projects which are not of strategic importance.

In order to successfully apply active resource allocation, the company's strategy has to be clearly understood throughout all levels of the company. This forms the basis of actively engaging employees in the project management process. If the employees understand the company strategy, they are can also evaluate the projects in which they are involved on the basis of their importance for achieving the strategic goals (Bruch and Menges, 2010).

A systematic process of regularly assessing the company's activities can be advisable and should consist of two dimensions: some of the control-processes should be aimed at assessing the ongoing activities, while other processes are meant to reduce the number of activities being started (Bruch and Vogel, 2011).

An inverse idea management system is an example of how ongoing activities can be assessed: Regularly ask employees which initiatives they think could be terminated. Another way is to regularly check which of the company's current activities would actually have been initiated if one had already known their present outcomes – those that get a negative rating in this procedure should clearly be eliminated.[14]

Capping the number of goals each manager is allowed to pursue each year and filtering projects before they are started according to their strategic relevance are examples of strategies to limit the number of activities right from the beginning. Clear project goals, binding timetables and milestones, as well as a systematic staffing process and knowledge management ensure that those projects that are successful continue and that processes are also carried out efficiently and do not tie up too many resources.

It is important to emphasize that the success rate of establishing processes to control for project effectiveness and efficiency in order to free up resources is highly dependent on the company's culture (Bruch and Menges, 2010; Bruch and Vogel, 2011). In companies where terminating projects because they do not deliver the expected outcome is not valued, but the project team's careers could instead suffer negative consequences, or where unfocussed business is valued higher than focused decision making, an active cultural change is a necessary first step to stop wasting resources.

Currently, research on the acceleration trap is also unveiling structural antecedents of over-acceleration; these are factors that tend to consume resources in times of normal workload but intensify the acceleration trap's negative effects in times of growth. At this time, two major groups of factors have been identified: factors that hinder employees from making quick decisions by limiting their field of influence and factors that lead to internal conflicts which undermine organizational units' willingness to cooperate (Kunz, 2011).

Executives should therefore avoid high levels of centralization whenever possible and thus, allowing their employees to make quick decisions. They should also control for overlapping individual goals and unequally distributed resources (e.g., information) that foster conflicts between employees, teams, or even entire com-

[14] For a detailed description on how to establish inverse idea management systems, see Bruch and Vogel (2011).

pany units. Not being sensitive to these factors means wasting resources on slow decision processes and internal conflicts that, especially in times of growth and high demands, are needed elsewhere in the organization.

By monitoring for over-acceleration and installing processes to systematically fight the antecedents of the acceleration trap, companies can control for quantitative overloading that can endanger balanced growth. Nevertheless, it is equally important to make sure that employees have the necessary skills to handle organizational growth and do not experience qualitative overloading.

4.3 Lack of key skills

From an HR perspective, the precondition for every form of successful organizational growth is a certain set of skills that the personnel has to possess (Ulrich, 1996). Not only does growth lead to more complexity (Penrose, 1959), and therefore requires personnel skilled in coping with the subsequent increase in demands, but organic growth (Hagel, 2002), as well as growth via acquisitions (Christensen et al., 2002), is also only possible if the personnel has the required skills to identify and exploit growth opportunities. In the case of organic growth (Hagel, 2002), this means the need for skills to identify customer needs in the context of existing business relationships in order to initiate growth by leveraging customers and/or to turn the company's core competencies into product innovations, thereby entering new areas of business (Ulrich, 1996). In the case of growth via acquisitions, besides the need for skills to identify potential acquisition candidates' financial and strategic fit, there is also the often neglected need to foster skills that enable the personnel to deal with the cultural aspects of merging different corporate cultures (Ulrich, 1996).

This finding has a clear implication for companies relying on their existing personnel to facilitate sustainable growth: It is not enough to only ensure that they master the quantitative HR challenges of growth, it is also of immense importance that the skills necessary to identify and exploit growth opportunities exist within the company. This is the qualitative challenge of growth.

According to the 2010 Accenture High Performance Workforce Study, which captured the opinions of 674 executives, the probability is small that most firms' personnel currently possess the necessary skills to master this qualitative challenge of sustainable growth (Farley and Gartside, 2011). Some of the study's key statistics summarize the problem: Only a minority of the participating executives rated their staff as high performing (21 %), while the vast majority of them admitted that their staff's general skills level left considerable room for improvement (84%)[15]. The true extent of this problem, the gap between personnel skills and the requirements of the firm's surrounding, becomes obvious when analyzing the executive's answers concerning their companies' ability to quickly exploit new growth opportunities (e.g., the execution of new strategies, entering new market, product innovation, etc.). Only one third of the respondents strongly believed their personnel capable of

[15] The reported results are derived from two different questions used in the study; consequently, the results do not add up to 100%.

doing so. These sobering results underline that the necessity to strategically monitor and manage the personnel skill level in order to ensure sustainable growth is a challenge that a majority of companies do not handle adequately. What is the reason for this blind spot in managing growth?

The Accenture study identifies the capabilities of the HR departments as the key bottleneck in supporting the growth of companies. 92% of the interviewed executives pointed out that they were not fully satisfied with their HR department's performance in supporting the execution of the company's strategic goals. Farley and Gartside (2011) therefore interpret their results as a wakeup call to enhance the alignment between company strategy and HRM, which emphasizes HRM's strategic importance. The authors conclude: "The road back to […] growth begins with companies aligning their business strategy with a human capital strategy that puts in place the right talent in the right roles performing in the right ways to bring the business strategy to life an execute it optimally" (p. 48).

Farley and Gartside (2011) are not the first to highlight the need for a strategic HRM. According to Wright et al. (2001), the increasing emphasis of HRM's strategic importance is closely linked to the introduction of the resource-based view of the firm. This paradigmatic shift directed the focus of research from competitive advantages to companies' internal resources (Hoskisson et al., 1999). When this new focus was applied, it became obvious that the company personnel also had to be regarded as a strategic resource (Wright et al., 2001). Teece et al. (1997) summarize the novel insights into the role of HRM gained from this paradigm shift: "Indeed if control over scarce resources is the source of economic profits, then it follows that such issues as skill acquisition, the management of knowledge and know how and learning become fundamental strategic issues. It is in this second dimension, encompassing skill acquisition, learning and accumulation of organizational and intangible or invisible assets, that we believe lies the greatest potential for contributions to strategy" (pp. 514-515).

4.4 Interlink strategic HRM and high performance work systems

Ulrich (1996) proposes a two-step approach to deriving practical recommendations for day-to-day HRM from this strategic HRM perspective: Initially, companies need to identify their individual capabilities, which Ulrich (1996) defines as "the things an organization always does better than its competitors do" (p. 10). Regardless of whether these capabilities are hard (e.g., technological advantages) or soft factors (e.g., being highly innovative), the second step is to align the HRM activities to these capabilities. The goal is to make sure that the skills required to maintain and extend the competitive advantages derived from these capabilities are embedded in the personnel's individual competence sets.

A study by the Kienbaum consultancy involving the HR departments of 161 German, Austrian, and Swiss companies, revealed that many companies do not yet apply a sufficient strategic HRM (Kutasi et al., 2008). Over 30% of the analyzed companies do not align their HRM with the company's strategy, and 40% of the companies do not have an explicit HR strategy. Even more alarming, 54% of the

companies do not use a company-wide competence model that constitutes the basis for a systematic personnel skill management. Furthermore, it is striking that the study reveals an average transfer rate of less than 50% when it comes to evaluating whether skills acquired in trainings are observable in employees' day-to-day work behavior.

As far as the suboptimal transfer rate is concerned, it is important to keep in mind that Salaman et al. (2005) point out that "individuals as cognitive and emotional beings […] possess free will. This free will enables them to make decisions regarding the behaviors in which they will engage" (p. 21). From a strategic HRM perspective, this seemingly trivial observation has profound implications concerning the acquisition of skills. Simply stated, it means that in order to maximize the company's competitive advantage, training the personnel's skills is not sufficient for a successful strategic HRM. There are also contextual factors which determine whether the staff members are able and, most importantly, willing to apply the learned skills. Salaman et al. (2005) come to the conclusion: "Firms may have access to valuable human capital, but either through the poor design of work or the mismanagement of people, may not adequately deploy it to achieve strategic impact" (p. 21).

How can companies then fully exploit the potential advantages of strategic HRM? From a scientific point of view, high performance work systems (HPWS), which are also known as high-involvement or high-performance HR systems, seem to provide a promising framework for a successful strategic HRM, as they have empirically been proven to have a positive performance link (Becker and Gerhart, 1996; Huselid, 1995; Ichniowski et al., 1996). The majority of authors characterize HPWS as generally consisting of "coherent sets of HR practices that enhance employee skills, participation in decisions, and motivation" (Batt, 2002, p. 587).

According to Huselid (1995), HPWS cover three different dimensions of HRM to ensure the goals defined by Batt (2002) are met. First, HPWS apply HRM practices to attain highly skilled employees. This dimension includes the use of recruiting procedures which not only help build a sufficient pool of highly skilled potential candidates, but also allow choosing employees who best fit the profile of the vacant job from this pool. This first dimension also includes all HRM instruments that can ensure that the skill level or the employee's development is continuously improved, for example, by providing adequate trainings, coaching, etc. (Huselid, 1995).

Based on the notion that even if the first dimension is implemented properly, and the company therefore possesses highly skilled personnel, the absence of motivation to perform will render all these effort useless, Huselid (1995) defines employee motivation as the second HPWS dimension. This dimension can utilize "performance appraisals that assess individual or work group performance," which are linked "with incentive compensation systems," internal promotion systems that focus on employee merit, as well as "other forms of incentives intended to align the interests of employees with those of shareholders" (Huselid, 1995, p. 637).

The third HPWS dimension utilizes organizational structures to create an environment that fosters opportunities for employees to use their skills, as well as their abilities. Furthermore, such an environment also encourages "participation among

employees" and allows them "to improve how their jobs are performed" (Huselid, 1995, p. 638). This dimension includes practices like "cross-functional teams, job rotation, and quality circles" (Huselid, 1995, p. 638).

Consequently, implementing HPWS as a framework for strategic HRM can be deemed a promising strategy to ensure that employees really apply those skills that they have acquired during trainings in the day-to-day business context. If the trainings content is derived from a capability-focused HR strategy, chances are good that the company could master the qualitative challenges of sustainable growth.

5 Establish a sustainable, balanced growth

As soon as shareholder interests come into play, growth currently seems to be an inevitable imperative for publicly traded companies specifically, but also for privately owned companies (Christensen et al., 2003). But this is only half the story, as Christensen et al. (2003) demonstrate that the equity market also expects sustainable, and therefore long-lived, growth. Empirical data shows that shareholders severely punish those companies experiencing stalling growth rates.[16]

Growth – regardless of whether it is achieved by acquiring other firms (e.g., Christensen et al., 2011) or by organic expansion (e.g., Hagel, 2002) – always leads to more complexity (Penrose, 1959). Consequently, sustainable growth demands a continual increase in resources. In the context of the irreversible shift from the industrial age to the information age (Michaels et al., 2001) and the increasing complexity of job profiles, it is currently often human capital resources which are most urgently needed to master growth. This is bad news, as the talent pool of young and well-educated workers is shrinking in many countries (Michaels et al., 2001). Thus, the resulting "war for talent" and the management of knowledge, know-how, and learning become fundamental strategic issues for companies wanting to gain a significant competitive advantage.

Since the management and development of human resources are primary areas of action for HR departments, this chapter presented HRM strategies that empower companies to manage the rising demand for human capital resources successfully in order to ensure sustainable growth. Two general approaches have been identified: First, acquiring new personnel (as described in Section 3) and, second, optimizing the employment of existing staff (as described in Section 4).

However, at this point it is important to stress that these two strategies are not mutually exclusive. On the contrary, given the substantial consequences that the demographic change will have for the job market, interlocking both strategies will become increasingly important for companies. In other words, the process of establishing the own company as an attractive employer to appeal to those scarce, highly qualified job candidates (as described in Section 3.1) benefits from highly moti-

[16] In an empirical study reported by Christensen (2003), 69% of the studied 172 companies from the Fortune list reported that their market value dropped by at least 50% as soon as their growth rates stalled.

vated and committed personnel who act as a testimonial for the employer brand (Schuhmacher and Geschwill, 2008).

Avoiding over-acceleration and establishing a high performance work system (HPWS) framework in this respect has multiple benefits for companies. First of all, being sensitive to the risk of falling victim to the acceleration trap and fighting its antecedents lead to a more effective and efficient resource allocation (as described in Section 4.2). This not only significantly lowers the risk of organizational burnout, but also improves the organizational climate, as it reduces tensions and aggressive behavior, which in turn positively influences employee motivation. On the other hand, the HPWS framework not only significantly increases the probability that employees will use their skills to facilitate sustainable growth (as described in Section 4.4), but it also adds to work motivation, thereby increases commitment.

In the same vein, applying a strategic HRM supports both the HRM strategies to facilitate sustainable growth. Not only does strategic HRM allow for a more systematic assessment of the skills that the existing workforce needs to master the qualitative challenge of organizational growth (as described in Section 4.4), but it can also professionalize the recruiting process. By defining a clear set of organizational capabilities that act as framework for a company's HRM processes, HR professionals also can derive more detailed profiles for suitable job candidates. By precisely defining the necessary skills, as well as the values that a potential new employee requires to add to the organizational capabilities, strategic HRM becomes another strategy to avoid the dilution of the organizational culture/identity (as described in Section 3.4).

Finally, it should be noted that in order to fight the talent bench shortfall, changing the target group by recruiting older, highly skilled employees (as described in Section 3.2) can also be matched by strategic HRM practice measures. At present, there are still tendencies to exclude older employees from trainings as the return on the investment is not deemed high enough (e.g., Schröder et al., 2004). Given the consequences of the demographic change, it will become more and more important not to waste HR by denying older employees the possibility to keep their competences up to date.

Although the talent bench shortfall intensifies the struggle for sustainable and balanced growth, the outlined solutions can support companies in mastering this challenge. However, this requires a dramatic shift in the perception and the strategic integration of the HR-department, which many companies have yet to undertake (McKinsey, 2011). It is no longer sufficient to view HRM as an administrative task; HRM is transforming into a key strategic task in the "war" of winning and retaining human capital resources (McKinsey, 2011). Ulrich's (1996) conclusion concerning the changing role of HRM, which he formulated 15 years ago, can be seen as a mission statement for the upcoming transformation of the HRM-scope:

"HR departments are not designed to provide corporate therapy or as social or health-happiness retreats. HR professional must create the practices that make employees more competitive, not more comfortable. […] HR practices must create value by increasing the intellectual capital within the firm" (p. 18).

References

Allen, T. D., & Eby, L. T. (2010). *The Blackwell handbook of mentoring: A multiple perspectives approach*. Oxford, UK: Wiley-Blackwell.

Backhaus, K., & Tikoo, S. (2004). Conceptualizing and researching employer branding. *Career Development International*, 9(5), 501-517.

Badaracco, J. L. (1998). The discipline of building character. *Harvard Business Review*, 76, 114-125.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1986). *Social foundations of thought & action*. Englewood Cliffs, NJ: Prentice-Hall.

Barney, J. B. (1986). Organizational culture: Can it be a source of sustained competitive advantage? *Academy of Management Review*, 11(3), 656-665.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.

Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology*, 94(6), 1394-1411.

Bass, B. M. (1990). From transactional to transformational leadership: Learning to share the vision. *Organizational Dynamics*, 18(3), 19-31.

Bass, B. M., & Avolio, B. J. (1993). Transformational leadership and organizational culture. *Public Administration Quarterly*, 17, 112-121.

Bass, B. M., Waldman, D. A., Avolio, B. J., & Webb, M. (1987). Transformational leadership and the falling dominoes effect. *Group and Organization Studies*, 12, 73-87.

Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal*, 45(3), 587-597.

Becker, B., & Gerhart, B. (1996). The Impact of human resource management on organizational performance: Progress and prospects. *Academy of Management Journal*, 39(4): 779-801.

Beechler, S., & Woodward, I. C. (2009). The global 'war for talent'. *Journal of International Management*, 15(3), 273-285.

Bellman, E. & Kistler, W. J. (2003). Betriebliche Sicht- und Verhaltensweisen gegenueber aelteren Arbeitnehmern. *Aus Politik und Zeitgeschichte*, 20, 26-34.

Berthon, P., Ewing, M., & Hah, L. (2005). Captivating company: dimensions of attractiveness in employer branding. *International Journal of Advertising*, 24(2), 151-172.

BKK (2011). BKK Gesundheitsreport 2010 – Gesundheit in einer älter werdenden Gesellschaft. Resource document. BKK Bundesverband. http://www.bkk.de/fileadmin/user_upload/PDF/Arbeitgeber/gesundheitsreport/BKK_Gesundheitsreport_2010.pdf. Accessed 7 July 2011.

Boehm, S. A., Kunisch, S., & Boppel, M. (2011). An integrated framework for investigating the challenges and opportunities of demographic change. In S. Kunisch, S. A. Boehm & M. Boppel (Eds.), *From Grey to Silver* (pp. 3-21). Berlin Heidelberg: Springer.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, 34(6), 1080-1109.

Bowers, C. A., Pharmer, J. A., & Salas, E. (2000). When member homogeneity is needed in work teams. *Small Group Research*, 31(3), 305-327.

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *Leadership Quarterly*, 17, 595-616.

Bruch, H., & Menges, J. I. (2010). The Acceleration Trap. *Harvard Business Review*, 88(3): 80-86.

Bruch, H., & Vogel, B. (2011). *Fully Charged – How Great Leaders Boost Their Organization's Energy and Ignite High Performance*. Boston: Harvard Business School Press.

Chao, G. T., Walz, P., & Gardner, P. D. (1992). Formal and informal mentorships: A comparison on mentoring functions and contrast with nonmentored counterparts. *Personnel Psychology*, 45(3), 619-636.

Christensen, C., M., Alton, R., Rising, C., & Waldeck, A. (2011). The new M&A playbook. *Harvard Business Review*, 89(3), 48-57.

Christensen, C. M., & Raynor, M. E. (2003). *The innovators solution: Creating and sustaining successful growth*. Boston: Harvard Business School Press.

Cole, M. S., & Bruch, H. (2006). Organizational identity strength, identification, and commitment and their relationships to turnover intention: does organizational hierarchy matter? *Journal of Organizational Behavior*, 27(5), 585-605.

Collins, J. C., & Porras, J. I. (1996). Building your company's vision. *Harvard Business Review*, 74, 65-78.

Davis, A., & Rothstein, H. (2006). The effects of the perceived behavioral integrity of managers on employee attitudes: A meta-analysis. *Journal of Business Ethics*, 67, 407-419.

DeLong, T. J., Gabarro, J. J., & Lees, R. J. (2008). Why mentoring matters in a hypercompetitive world. *Harvard Business Review*, 86(1), 115-121.

Deutsches Statistisches Bundesamt (2009). Bevölkerung Deutschlands bis 2060. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pk/2009/Bevoelkerung/pressebroschuere__bevoelkerungsentwicklung2009,property=file.pdf. Accessed 05 June 2011.

Deutsches Statistisches Bundesamt (2010a). Wanderungen von und nach Deutschland. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Zeitreihen/LangeReihen/Bevoelkerung/Content75/lrbev07a,templateId=renderPrint.psml. Accessed 05 June 2011.

Deutsches Statistisches Bundesamt (2010b). Bevölkerung nach dem Gebietsstand. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Zeitreihen/LangeReihen/Bevoelkerung/Content75/lrbev03a,templateId=renderPrint.psml. Accessed 05 June 2011.

Farley, C., & Gartside, D. (2011). Talent that fuels growth. *Talent Management Magazin*, 6(8): 42-48.

Farrell, D., & Grant, A., (2005). China's looming talent shortage. *McKinsey Quarterly*, 4, 70-79.

Frook, J. E. (2001). Burnish your brand from the inside. *B to B*, 86(8), 1-2.

Gioia, D. A., & Thomas, J. B. (1996). Identity, image, and issue interpretation: Sensemaking during strategic change in academia. *Administrative Science Quarterly*, 41(3), 370-403.

Glebbeek, A. C., & Bax, E. H. (2004). Is high employee turnover really harmful? An empirical test using company records. *Academy of Management Journal*, 47(2), 277-286.

Hagel, J. (2002). Leveraged growth: Expanding sales without sacrificing profits. *Harvard Business Review*, 80(10), 68-77.

Hoskisson, R. E., Hitt, M. A., Wan, W. P., & Yiu, D. (1999). Theory and research in strategic management: Swings of a pendulum. *Strategic Management Journal*, 25 (3): 417-456.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38(3): 635-672.

Ichniowski, C., Kochan, T., Levine, D., Olson, C., & Strauss, G. (1996). What works at work: Overview and assessment. *Industrial Relations*, 35(3): 299-334.

Jacobi, M. (1991). Mentoring and undergraduate academic success: A literature review. *Review of Educational Research*, 61(4), 505-532.

Kacmar, K. M., Andrews, M. C., Rooy, D. L. V. A. N., Steilberg, R. C., & Cerrone, S. (2006). Sure everyone can be replaced... but at what cost? Turnover as a predictor of unit-level performance. *Academy of Management Journal*, 49(1), 133-144.

Kunz, J. J. (2011). Die Beschleunigungsfalle: Aktuelle empirische Erkenntnisse 2011. Speech presented at the St.Gallen Leadership-Day 2011, St.Gallen.

Kutasi, C., Lassalle, J., El Ganady, Y., König, C., & Urban, W. (2008). Kienbaum Studie 2008 Personalentwicklung. Resource document. Kienbaum Consultants International GmbH. http://www.kienbaum.de/ext/download/kienbaum_studie_2008_personalentwicklung.pdf. Accessed 7 July 2011.

Lencioni, P. M. (2002). Make your values mean something. *Harvard Business Review*, 80(7), 113-117.

Lippman, S. A., & Rumelt, R. P. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *The Bell Journal of Economics*, 13(2), 418-438.

Lockett, A., Thompson, S., & Morgenstern, U. (2009). The development of the resource-based view of the firm: A critical appraisal. *International Journal of Management Reviews*, 11, 9-28.

Loretto, W., & White, P. (2006). Employers' attitudes, practices and policies towards older workers. *Human Resource Management Journal*, 16(3), 313-330.

Malmberg, B., Lindh, T., & Halvarsson, M. (2008). Productivity consequences of workforce aging: Stagnation or Horndal effect? *Population and Development Review*, 34, 238-256.

McElroy, J. C., Morrow, P. C., & Rude, S. N. (2001). Turnover and organizational performance: A comparative analysis of the effects of voluntary, involuntary, and reduction-in-force turnover. *Journal of Applied Psychology*, 86(6), 1294-1299.

McEvoy, G. M., & Cascio, W. F. (1989). Cumulative evidence of the relationship between employee age and job performance. *Journal of Applied Psychology*, 74(1), 11.

McKinsey. (2011). Wettbewerbsfaktor Fachkräfte – Strategien für Deutschlands Unternehmen. Online document. McKinsey. http://www.mckinsey.de/downloads/presse/2011/wettbewerbsfaktor_fachkaefte.pdf. Accessed 04 July 2011.

Meinert, S. (2011). Burn-out und andere teure Probleme. Resource document. Financial Times Deutschland. http://www.ftd.de/karriere-management/karriere/:mitarbeitergesundheit-burn-out-und-andere-teure-probleme/60062315.html. Accessed 7 July 2011.

Michaels, E., Handfield-Jones, H., & Axelrod, B. (2001). *The war for talent*. Boston, MA: Harvard Business Press.

Ng, T. W. H., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology*, 93(2), 392-423.

Ng, T. W. H., Sorensen, K. L., Eby, L. T., & Feldman, D. C. (2007). Determinants of job mobility: A theoretical integration and extension. *Journal of Occupational and Organizational Psychology*, 80(3), 363-386.

Olson, M. S., van Bever, D., & Verry, S. (2008). When growth stalls. *Harvard Business Review*, 86(3), 50-61.

Kanfer, R., & Ackerman, P. L. (2004). Aging, adult development, and work motivation. *Academy of Management Review*, 29(3), 440-458.

Kroth, M., & Keeler, C. (2009). Caring as a managerial strategy. *Human Resource Development Review*, 8(4), 506-531.

Penrose, E. T. (1959). *The Growth of the Firm*. New York: Wiley.

Pfeffer, J. (1994). *Competitive advantage through people – unleashing the power of the workforce*. Boston, MA: Harvard Business School Press.

PricewaterhouseCoopers (2011). The 14th annual global CEO survey. http://www.pwc.com/gx/en/ceo-survey/download.jhtml. Accessed 31 May 2011.

Ray, S. C., & Jeon, Y. (2008). Reputation and efficiency: A non-parametric assessment of America's top-rated MBA programs. *European Journal of Operational Research*, 189(1), 245-268.

Salaman, G., Storey, J., & Billsberry, J. (2005). *Strategic Human Resource Management: Theory and Practice*. London: Sage.

Schein, E. H. (1983). The role of the founder in creating organizational culture. *Organizational Dynamics*, 12(1), 13-28.

Schein, E. H. (2010). *Organizational culture and leadership* (4 ed.). San Francisco, CA: Jossey-Bass.

Schneider, B. (1983). An interactionist perspective on organizational effectiveness. In L. L.Cummings & B. M.Staw (Eds.), *Research in organizational behavior* (pp. 1-31). Greenwich, CT: JAI Press.

Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40(3),
437-453.

Schröder, H., Schiel, S., & Aust, F. (2004). *Nichtteilnahme an beruflicher Weiterbildung. Motive, Beweggründe, Hindernisse*. Bielefeld: Expertenkommission Finanzierung Lebenslangen Lernens.

Schuhmacher, F., & Geschwill, R. (2009). *Employer Branding: Human Resources Management für die Unternehmensführung*. Wiesbaden: Gabler Verlag.

Schuler, R. S., & Jackson, S. E. (2007). *Strategic human resource management* (2nd ed.). Oxford: Blackwell.

Schweizer Bundesamt für Statistik (2009). Demografisches Porträt der Schweiz 2009. http://www.bfs.admin.ch/bfs/portal/de/index/news/publikationen.Document.127584.pdf. Accessed 20 July 2010.

Shaw, J. D., Gupta, N., & Delery, J. E. (2005). Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. *Academy of Management Journal*, 48(1), 50-68.

Sheridan, J. E. (1992). Organizational culture and employee retention. *Academy of Management Journal*, 35(5), 1036-1056.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7): 509-533.

Ulrich, D. (1996). *Human resource champions*. Boston: Harvard Business School Press.

Van Dalen, H. P., Henkens, K., & Schippers, J. (2010). Productivity of older workers: perceptions of employers and employees. *Population and Development Review*, 36(2), 309-330.

Van Hoye, G., & Saks, A. M. (2011). The instrumental-symbolic framework: Organizational image and attractiveness of potential applicants and their companions at a job fair. *Applied Psychology – An International Review*, 60(2), 311-335.

Van Ours, J. C., & Stoeldraijer, L. (2010). Age, wage and productivity. Working paper. CESifo, Center for Economic Studies & Ifo Institute for economic research. Tilburg, NL.
http://www.ifo dresden.de/portal/page/portal/DocBase_Content/WP/WP-CESifo_Working_Papers/wp-cesifo-2010/wp-cesifo-2010-02/cesifo1_wp2965.pdf. Accessed 09 June 2011.

Van Yperen, N. W., & Hagedoorn, M. (2003). Do high job demands increase intrinsic motivation or fatigue or both? The role of job control and job social support. *Academy of Management Journal*, 46(3), 339-348.

Waldman, D. A., & Avolio, B. J. (1986). A meta-analysis of age differences in job performance. *Journal of Applied Psychology*, 71(1), 33-38.

Wegge, J., Roth, C., Neubach, B., Schmidt, K. H., & Kanfer, R. (2008). Age and gender diversity as determinants of performance and health in a public organization: The role of task complexity and group size. *Journal of Applied Psychology*, 93(6), 1301-1313.

Wey Smola, K., & Sutton, C. D. (2002). Generational differences: Revisiting generational work values for the new millennium. *Journal of Organizational Behavior*, 23(4), 363-382.

Wilson, J. A., & Elman, N. S. (1990). Organizational benefits of mentoring. *Academy of Management Executive*, 4(4), 88-94.

World Economic Forum. (2011). Global talent risk – seven responses. Online document. World Economic Forum. http://www.bcg.de/documents/file69643.pdf. Accessed 03 July 2011.

Wright, P. M., Dunford, B. B., & Snell, S. A. (2001). Human resources and the resource based view of the firm. *Journal of Management*, 27(6): 701-721.

Zachary, L. J. (2005). *Creating a mentoring culture*. San Francisco, CA: Jossey-Bass.

Zaugg, R. J., & Thom, N. (2003). Excellence through Implicit Competences. Human Resource Management – Organizational Development – Knowledge Creation. *Journal of Change Management*, 3(3): 199-211.

Zott, C. (2003). Dynamic capabilities and the emergence of intraindustry differential firm performance: Insights from a simulation study. *Strategic Management Journal*, 24(2), 97-125.

Electricity storage: Making large-scale adoption of wind and solar energies a reality

Cornelius Pieper and Holger Rubel

Abstract:

The growth case for using renewable energy remains very much intact, despite the effects of a global economic downturn, and the prospects for wind and solar photovoltaic (PV) power appear particularly strong. To realize this potential, however, these technologies will have to overcome a key hurdle: the challenge posed by their intermittent nature. Unlike other forms of renewable energy, such as hydropower and geothermal energy, the energy generated by wind and solar PV fluctuates. This fluctuation poses a sizable challenge to their integration into the power grid and their widespread adoption as bona fide mainstream power sources.[1] While there are several potential answers to the challenges of intermittency, the most viable, we believe, is a credible form of electricity storage. Yet there is little on the immediate horizon to suggest that a storage solution is imminent. Unless and until that vital enabler exists and is practical, fluctuating renewables will struggle to become key players in the global push toward carbon dioxide–free energy sources. This study is based on extensive research and more than 30 interviews conducted by the Boston Consulting Group (BCG) with industry experts in late 2009, and considers the current state of play in electricity storage. Although we take a global perspective, we focus particularly closely on Europe, which will be the first region to experience large-scale storage challenges. We examine why the need for storage will only grow, what the key technologies are, where they are developmentally, and what their prospects are for adoption; and what the potential alternatives to storage are, and why none is sufficient on its own. We also discuss implications for stakeholders. Key findings include:

- Electricity storage will be essential for successful, critical mass adoption of fluctuating renewables. Although alternative solutions such as interregional compensation, demand-side management, and conventional backup power exist and will continue to play important roles, the extent and degree of fluc-

[1] We refer to wind and solar PV as "fluctuating renewables" throughout this paper. Furthermore, we use the term to refer to the electricity generated from fluctuating renewables. We also distinguish between solar PV and concentrated solar power (CSP), the other main solar technology. In many cases, CSP has an internal storage capability to compensate for fluctuations. Thus, it already feeds a "flattened" power curve into the grid and does *not* typically require external grid storage. Hence, in this paper, we confine our discussion of solar energy to solar PV.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_11,

tuation resulting from increased deployment of fluctuating renewables will demand the use of storage technologies.

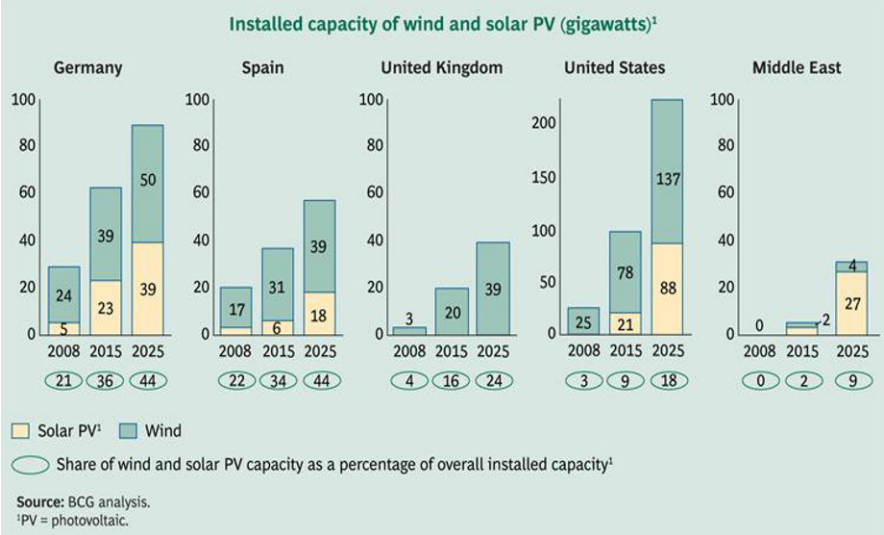
- Because the financial logic for investing aggressively to advance storage technologies is currently not compelling, incentives will be necessary to ensure that sufficient storage capacity is online in time to meet governments' green energy targets.
- Players that actively participate in shaping the technological, political, and market frameworks that determine the future use of storage technologies stand to gain a clear competitive edge, since technologies will be chosen and established, policies will be negotiated and deployed, and R&D partnerships and customer relationships will be established and strengthened in the coming years.

Table of Contents

1 The growth of wind and solar PV energies means a growing need for compensating capacity

The need for compensating capacity will rise in lockstep with the growth in demand for fluctuating renewables. The demand for power, gauged by the growth of installed capacity, stands to rise strongly in several global regions through 2025, reflecting governments' intensifying efforts to move away from fossil-based fuels. (See Exhibit 1.)

Exhibit 1. Growth in Demand for Wind and Solar PV Energies Will Likely Be Strong Across All Major Regions Through 2025

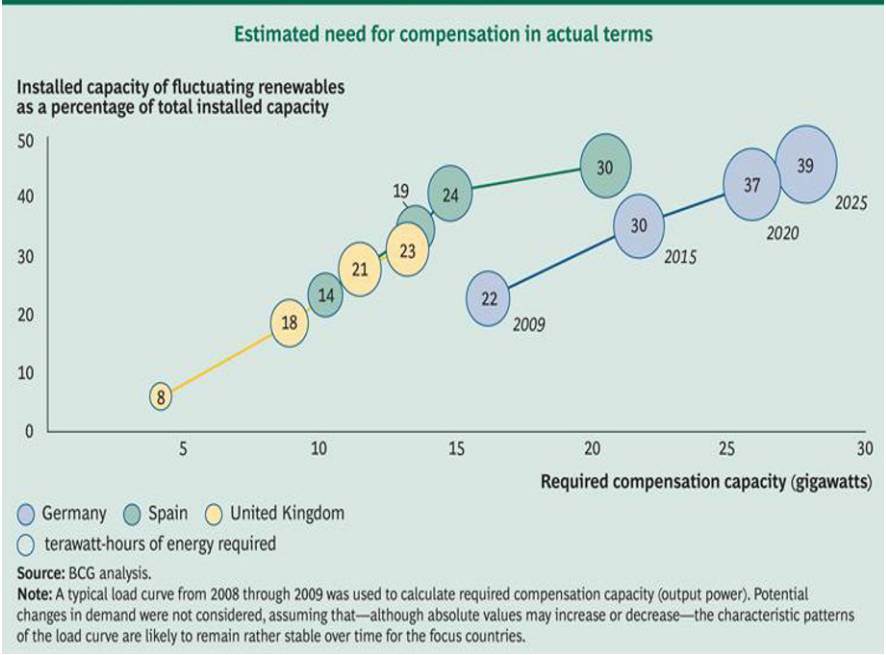


Source: BCG analysis.
[1]PV = photovoltaic.

On balance, the greater the share of fluctuating renewables in a given energy-generation region, the greater the fluctuations produced. Hence, the greater the need for compensating capacity. (See the sidebar 1.) But the extent of fluctuation will also be influenced by the mix of wind and solar PV energies within a particular system: fluctuations in one source can be partially offset by fluctuations in the other.

In Exhibit 2, we show demand projections for compensation capacity resulting from rising penetration of fluctuating renewables in Germany, Spain, and the UK for four reference years. As Europe will be the first region to experience problems related to fluctuating renewables, we have focused on Europe for quantification of the challenge. In principle, however, our findings will also hold true for other regions.

Exhibit 2. The Need for Compensating Capacity Will Rise Sharply in Germany, Spain, and the United Kingdom Through 2025



Source: BCG analysis.
Note: A typical load curve from 2008 through 2009 was used to calculate required compensation capacity (output power). Potential changes in demand were not considered, assuming that—although absolute values may increase or decrease—the characteristic patterns of the load curve are likely to remain rather stable over time for the focus countries.

These estimates are based on what we consider realistic assumptions regarding compensation demands imposed by both predictable (i.e. day-night and seasonal) and unpredictable (i.e. driven by medium-term weather conditions and forecast errors) variability in fluctuating renewables. In 2025, as much as 28 gigawatts of compensating capacity will be required in Germany to provide up to 40 terawatt hours of compensation energy. Spain and the UK will have similar requirements. In contrast, current storage capacities stand at about 7 gigawatts in Germany, 5 gigawatts in Spain, and 4 gigawatts in the UK. Almost all of this existing storage is in the form of pumped hydroelectric storage facilities.

Collectively, European countries will need about 100 gigawatts of compensating capacity. This corresponds to a total installed generation capacity in 2025 of approximately 1,000 gigawatts, roughly 350 gigawatts of which will come from renewable sources.[2] The facilities will need to be capable of providing roughly 150 terawatt hours or more of compensation energy, corresponding to more than 5% of the annual demand for electricity.

The US, which has lower penetration of fluctuating renewables but a larger installed capacity in absolute terms, will need up to 170 gigawatts of compensating

[2] Eurostat.

capacity by 2025, considerably more than the roughly 25 gigawatts it has today. This corresponds to a forecast generation capacity of about 1,100 gigawatts in 2025, with renewable sources accounting for roughly 250 gigawatts.[3] Given the US's current difficulties in maintaining grid stability, the timely availability of adequate compensating capacity will be critical for a successful buildup of fluctuating renewables' capacity there.

Insufficient ability to compensate for fluctuations is not a far-into-the-future scenario. Already today, there are periods in which feed-in from renewables is higher than off-peak electricity demand in some regions, especially in countries where renewables represent a high share of overall generation capacity – for example, Denmark and Germany. In such periods, there is very high volatility in electricity prices. Furthermore, wholesale prices frequently turn negative, as happened several times in Germany in 2009, with record negative prices as low as 500 EUR per megawatt hour despite a powerful grid infrastructure capable of shifting large amounts of energy across Europe. And this is only the beginning of the impact of increasing fluctuations from renewables on the power grid.

The upshot of the above is that strongly rising demand for fluctuating renewables will necessitate a significant ramping up of compensating capacity in the years ahead. Next, we explore the four broad approaches to providing this capacity.

2 Four approaches to compensation

There are four methods of compensating for variability in electricity generation from fluctuating renewables: interregional compensation (often referred to as grid extension), conventional backup capacity, demand-side management, and large-scale electricity storage. Each has its strengths and limitations.

2.1 Interregional compensation (grid extension)

Interregional compensation, often referred to as grid extension (although, technically, grid extension is a means of achieving interregional compensation), involves linking electricity grids from different regions (each with its own generation and demand profile) and transferring power from one to the other to compensate for fluctuations: grid A in Region X, for example, imports power from Grid B in Region Y when Grid A experiences periods of low wind (or sun) and high demand for power, and exports it when it experiences high wind and low demand, provided the grids' profiles are complementary.[4]

[3] US Energy Information Administration.

[4] Interregional or cross-border compensation is distinct from the challenge of strengthening and eliminating bottlenecks in national grids. However, it should be emphasized that stronger national grids are critical to the task of integrating fluctuating renewables.

Grid extension holds promise as a solution to the fluctuation challenge. In Europe, some experts even assume the future presence of a continental "copper plate."[5] However, grid extension also has the following significant limitations:

It cannot mitigate all types of fluctuations. On a Europe-wide level, for example, it cannot compensate for day-night fluctuations.

There are political barriers to implementation. Building new transmission lines entails an often-lengthy approval process and almost always meets strong public resistance.

Weather and climate conditions are relatively similar across large parts of Central Europe. This means that long high-voltage direct-current (HVDC) lines between two distant regions are required, offering only point-to-point links, which still require local tie-ins to the national grids. Alternating current (AC) grid infrastructure at the landing points of HVDC bridges will need to be upgraded as well. Similarly, large-scale grid projects, such as a North Sea supergrid, will require powerful links with inland load centers.

There are efficiency losses in transmission. Normal AC high-voltage transmission lines lose up to 15% for every 1,000 kilometers traveled; HVDC cables lose about 3%.

The effectiveness of grid extension will decrease as deployment of fluctuating renewables rises across Europe and the fluctuation problem becomes commonplace. In our view, this is the strongest argument against interregional compensation. For Denmark, for example, a small country with an already fairly high deployment of fluctuating renewables, interregional compensation currently works well. However, current importers of the country's excess energy may no longer be interested tomorrow, as the latter may be struggling with their own overcapacities.

In short, interregional grid extension will likely make an important contribution towards mitigating electricity generation variability induced by fluctuating renewables, and selective links between grids should be instrumental in solving local challenges. However, grid extension is not a standalone solution for the long run.

2.2 Conventional backup power

Conventional backup power, especially when driven by combined-cycle gas turbine (CCGT) and pure-gas plants, is a mature, very flexible technology that could – at least theoretically – provide all the required compensating capacity. Conventional backup power also offers competitive economic parameters when renewables' penetration is still at low levels. Hence, in the coming years, we believe that much of the compensation capacity required to balance the variability in output due to fluctuating renewables will continue to be provided by flexible conventional power plants.

[5] The so-called copper plate refers to the assumption of an unrestricted power network across Europe.

However, as is the case with grid extension, the following factors will limit the construction of new flexible conventional power plants (beyond the mere replacement of outgoing plants).

Public and political resistance to dedicating new locations to conventional power plants will likely be strong at the local level. In addition, it will be difficult to argue, from an environmental perspective, that new construction of fossil fuel-based plants is beneficial and logical, given that the ultimate goal is eventually to achieve CO_2-free electricity generation.

Backup capacity (mainly gas-driven capacity) will extend countries' dependence on fossil fuel exporters.

The economic risk associated with investing in the construction of new conventional power plants is increasing as a result of three factors: possible increases in CO_2 permit prices, uncertainty regarding fuel prices, and the risk of low utilization.[6]

Conventional capacity can compensate for troughs – but not peaks – in fluctuating renewables' production. Hence, inefficient "throttling" of fluctuating renewables' output is still necessary, further reducing the utilization of generation assets (in this case, fluctuating renewables, which in some cases have been heavily subsidized) and driving up overall power generation costs.

There is also a location-related challenge to consider. Renewable generation will be both centralized (in the cases of offshore wind and large ground-mounted solar PV installations) and decentralized (in the cases of small-scale onshore wind farms and rooftop and small-scale commercial solar PV applications). To minimize further strain on the grid infrastructure, the necessary conventional backup capacity will need to be located close to these respective sources. In particular, conventional backup generation will require a strong decentralized component in regions where small-scale wind and solar PV are present, necessitating sophisticated steering mechanisms.

In light of these considerations, we do not believe conventional backup capacity will be sufficient on its own or sustainable as we move toward a renewables-dominated electricity system in the long term. Still, in our view, conventional backup capacity will be indispensable for achieving the integration of renewable energy sources into the current power system in the coming years.

2.3 Demand-side management

Demand-side management (DSM) refers to the postponing or advancing of suitable demand elements among industrial and residential customers so as to minimize overall demand during peak periods. Prime candidates for DSM on the industrial side are such energy-intensive industries as aluminum and chemicals, which have the potential to become more flexible in managing their electrical loads, given the

[6] The *Handelsblatt*, February 7, 2010, reported that in Germany alone, 7 large projects for coal-fired power plants were canceled in 2009 and 2010, largely as a result of local protests and worries over the projects' uncertain economics due to unclear utilization prospects associated with fluctuating renewables.

right incentives. Properly incentivized, many if not most households could also be convinced to shift at least some of the operation of their electricity-intensive equipment (for example, air conditioners and heating systems) to off-peak periods.

However, DSM has its limitations. Most loads can be deferred for a short period of time, but longer deferrals risk creating production disturbances. DSM also demands large investments, such as smart meters and customized billing systems. In addition to being costly, these devices could raise issues of privacy that would need to be resolved. Furthermore, DSM requires behavioral adaptations by customers and sufficient pricing flexibility to actually drive changes in customer thinking. A savings of only 10 cents, for example, is unlikely to convince many customers to do their laundry at 03:00 rather than at a more convenient time.

DSM is currently being used mainly in the US for peak shaving and reducing strain on grids. However, DSM is no panacea. Indeed, two unrelated studies in Germany and the US found that DSM offers a demand reduction potential of only approximately 2% of peak load. All told, although DSM clearly has potential and should be explored, we expect its ultimate contribution to the smoothing of wind and solar PV fluctuations to be limited.

2.4 Large-scale electricity storage

Large-scale electricity storage refers to harnessing excess power generated in times of abundant availability (or low demand, or both) and releasing it later into the grid when power generation is low (or demand is high, or both). Large-scale electricity storage offers structural advantages over both interregional compensation and compensation by backup capacity. Unlike interregional compensation, storage provides a self-sufficient solution for one specific region and hence is not affected by increases in penetration of fluctuating renewables across the board. And unlike compensation by backup capacity, storage can deal with troughs *and* peaks in fluctuating renewables' output. Additionally, because it can reduce fluctuations close to the various generation sources, large-scale electricity storage translates into less strain on grids.

The approach is not perfect, however. All electricity storage technologies are inefficient to a degree: a portion of the energy fed into the system cannot be discharged later on and is lost. With typical efficiency factors ranging between 45% (for hydrogen) and 80% (for batteries), efficiency is the key weakness of electricity storage technologies and, to a large extent, accounts for the currently unfavorable business case for them. As with conventional power plants, the business case for individual facilities is further affected by other factors, including the cost of charging electricity, capital expenditures, operating expenses, and number of cycles. The actual utilization of a storage facility is a complex function of many parameters, such as weather, load vs. demand, and grid constraints.

Large-scale electricity storage also raises other questions. Technologically, it is still relatively immature for large applications other than pumped hydroelectric storage, its implementation entails additional costs on top of generation costs (including capital and operating expenditures and costs related to the loss of elec-

tricity), and to date it has had limited operational testing at large scale. Yet, despite these concerns, we consider large-scale electricity storage a key means of addressing the challenge of compensation – one that will likely have to provide a significant share of the compensating capacity necessary in future power systems.

3 Storage technologies as compensation capacity: Characterization, applicability, and economic viability

Many different energy-storage technologies exist, but most are in the pilot phase of development. There are five main types of storage and key technologies within these categories:

1. *Mechanical* storage, including pumped hydroelectric storage, compressed air energy storage (CAES), and flywheel energy storage
2. *Thermal* storage, including hot water storage, molten-salt storage, and phase-change material storage
3. *Electrical* storage, including supercapacitors and superconducting magnets
4. *Electrochemical* storage, including flow and static batteries
5. *Chemical*, or hydrogen, storage.

Among these, the technologies that look most promising today for large-scale deployment are CAES, hydrogen storage, batteries, and pumped hydroelectric storage. All of these technologies are capable of storing significant amounts of energy, which is essential for balancing fluctuating renewables. Some of the other technologies, such as flywheels and supercapacitors, offer good performance in providing high capacity amounts for short periods of time, but they deliver only very limited amounts of energy and have very high self-discharge rates.[7] Currently, however, there are relatively few examples of these technologies in large-scale use. The most common are pumped hydroelectric facilities, of which there are approximately 300. There are only two CAES facilities in operation – one in Huntorf (Germany) and one in Alabama (US) – and there are a few pilot installations of large-scale batteries (mainly sodium sulfur [NaS] high-temperature batteries) in the US and Japan.

The four technologies in focus have different profiles in terms of developmental maturity, cost, and hurdles to implementation. These are summarized in Exhibit 3. We touch on some of the technologies' most noteworthy characteristics below.

[7] In the case of flywheels, for example, the rate of self-discharge is 100% per day.

Exhibit 3. A-CAES, Hydrogen Storage, Batteries, and Pumped Hydroelectric Storage Are the Most Relevant Large-Scale Technologies

	Technological maturity	Key application focus	Site limitations	Public concerns
A-CAES[1]	Partially mature	Mainly for large-scale centralized applications	Most developed countries have the necessary potential storage caverns; Japan and Spain have very few potential sites, however	Little concern expected
Hydrogen storage	Has yet to be demonstrated on a large scale	Generally, very flexible in terms of capacity; particularly suitable for decentralized applications	No specific geological requirements	Potential for safety concerns; however, reference projects are running safely
Stationary batteries	NaS is relatively mature; redox flow remains to be demonstrated on a large scale[2]	Generally, very flexible; particularly suitable for decentralized applications	No specific geological requirements	Few environmental concerns expected except for those related to the disposal of chemicals
Pumped hydroelectric storage	Mature	Very well suited for centralized applications; not yet implemented on a small scale	Europe, including Norway, has a limited number of potential sites left	Potential for environmental concerns, given the profound impact on landscapes

key disadvantage key advantage

Source: BCG analysis.
[1]A-CAES = adiabatic compressed air energy storage.
[2]NaS = sodium sulfur.

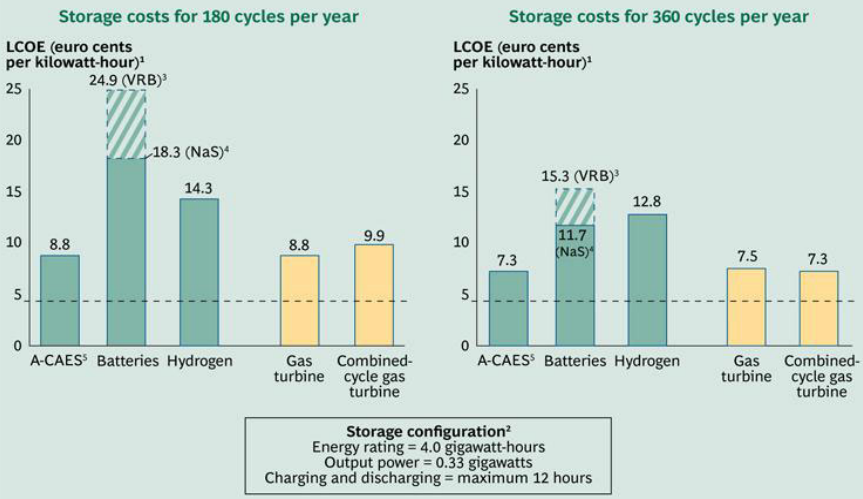
3.1 Compressed air energy storage

CAES, or the compression of air to be used later as an energy source is an existing technology that has proved its operational performance over many years in the abovementioned German and US installations. The working principle is simple: excess electricity is used to compress air into an underground cavern (or, for small installations, a surface tank). The pressurized air is subsequently released and used to drive the compressor of a natural-gas turbine, thereby creating electricity. In its current developmental state, CAES is relatively inefficient, with efficiency levels of only 45 to 55% – mainly the result of not retaining the heat obtained during compression. Also, if – when expanded – the compressed air were used to drive an air turbine, the turbine would basically freeze, because expanding air is cooling down. Since the decompressing air is used merely to support a conventional gas power plant, CAES in its current setup is not a self-sustaining energy storage method.

On the horizon, however, is so-called adiabatic compressed air energy storage (A-CAES), which we will now focus on. The main difference between A-CAES and conventional CAES is that A-CAES captures compression heat and stores it in a thermal storage unit. This heat, which is released upon discharging electricity, reheats the decompressing air. The heated air, in turn, drives an air turbine, which creates electricity. Hence, this technology is a self-sufficient storage solution. It remains in development, with its first pilot installations expected after 2013.

A-CAES is expected to be significantly more efficient than current CAES and should ultimately have a very attractive cost profile, compared with other forms of storage. (See Exhibit 4.) Indeed, our calculations indicate that A-CAES could be competitive on a cost basis with conventional forms of generation by 2025, provided its development continues as planned.[8]

Exhibit 4. By 2025, A-CAES Could Be Competitive with Conventional Generation



Source: BCG analysis.
Note: For conventional power plants, we assumed a gas price of €6 per million BTUs and a CO_2 permit price of €15 per ton.
[1]LCOE = levelized cost of energy.
[2]We assumed a cost of 4 euro cents per kilowatt-hour for charging energy. The LCOE for storage was calculated with estimated 2025 costs. Input power was selected accordingly to guarantee a 12-hour charging cycle.
[3]VRB = vanadium redox battery.
[4]NaS = sodium sulfur.
[5]A-CAES = adiabatic compressed air energy storage.

A-CAES will have to deal with a hurdle faced by all compressed-air storage technologies: there must be a sufficient number of suitable caverns or other geological formations for storage. This will pose little problem in certain locations, such as Northern Germany and large parts of the US, but it could be more problematic in other regions, including Spain and Japan, where suitable sites are rare.

3.2 Hydrogen storage

Hydrogen storage refers to using electricity to produce hydrogen with alkaline or proton exchange membrane electrolysis; storing the hydrogen, pressurized in an

[8] We assumed a rate of 4 EUR cents per kilowatt hour for the electricity used to charge the storage facility. But operators of storage facilities may achieve additional savings through lower grid fees resulting from more balanced grid utilization with lower power.

underground cavern (similar to compressed-air storage) or in a surface tank; and using the hydrogen for the generation of thermal electricity. Relative to A-CAES, hydrogen storage has somewhat higher investment costs and lower efficiency. Simultaneously, it has significantly higher energy density and hence significantly higher energy capacity. This, combined with its low efficiency, makes hydrogen storage most appropriate for the compensation of long-term (for example, seasonal) fluctuations.

Indeed, only hydrogen storage is capable of storing the huge amounts of electricity required for balancing seasonal fluctuations. In a seasonal operating mode with only one or two cycles per year, the disadvantages posed by hydrogen storage's low efficiency are also less of an issue. In fact, in the very same cavern, approximately 65 times more energy can be stored with hydrogen than with A-CAES. Because of this high energy efficiency, hydrogen storage may well emerge as the only storage option that can enable the balancing of seasonal fluctuations, which is a prerequisite for a 100% renewables-based electricity system. Implementing seasonal storage implies that storage infrastructure is set up to run only one or two cycles a year, which is likely to result in high storage costs per unit of energy.

Storage concepts that rely on hydrogen do not necessarily require costly infrastructure, as hydrogen is produced, stored, and consumed at the same site – or, alternatively, directly used for purposes other than power generation. However, a challenge facing hydrogen storage is whether it can operate efficiently at sufficiently large scale. Although the maturity of the technology, especially that portion required to produce hydrogen from electricity, has been demonstrated on smaller installations of up to 2 megawatts of power intake, it remains to be seen whether it can be scaled up to reach greater capacities.

Another challenge facing hydrogen storage is concern for public safety. Although the actual safety hazard posed by the technology, especially in the case of cavern storage, is relatively low, the industry may need to engage in educational efforts to overcome the public's reservations.

3.3 Stationary batteries

Stationary batteries, such as NaS batteries and vanadium redox batteries (VRBs), have the potential to become attractive choices for storage, especially in midsize applications, where cavern and turbine size limit other storage technologies. Stationary batteries are rechargeable. They are also the most efficient and flexible of the storage alternatives and are thus particularly cost competitive when the number of cycles is high, i.e. when used to compensate for day-night fluctuations. Battery storage also has relatively few site requirements, and there will likely be limited resistance to battery storage from the public.

It is noteworthy that electrolyte tanks and fuel cells are separate components of VRBs and all other flow batteries. This means that the energy size and power rating can be scaled independently, increasing the flexibility of these batteries.

Stationary batteries are currently the focus of much attention – especially in Japan, a leader in the development of battery technologies. Stationary batteries are also garnering attention in the US in the context of so-called community storage schemes, which are midsize decentralized storage facilities aimed at managing fluctuations from renewables and at reducing strain on the distribution grids.

The main hurdles for batteries are high investment costs, though these are expected to decrease over the next decade, driven by scale benefits and technology innovations. For both NaS batteries and VRBs, cost reductions of 30% or more in the next 10 years are possible, leading to investment costs of well below 2,000 EUR per kilowatt for an entire system. As most current battery installations come with average energy sizes for 6 to 8 hours of operation per cycle (1 megawatt of capacity translates into 6 to 8 megawatt hours of stored energy), the investment cost per kilowatt hour would be around 300 EUR.

The much-discussed idea of virtually bundling electric car batteries (utilizing so-called vehicle-to-grid technology) to create storage is appealing, but will not be realistic until the electric car significantly increases its market share. We do not expect that to happen before 2025. First, the absolute number and size of individual batteries will limit storage capacity. Second, the revenue potential for owners who make their batteries available to central dispatching will likely be too low to overcome owners' reluctance to shortening their battery's lifespan or having a partially discharged battery at an inopportune time. Third, the lithium ion batteries used in electric cars (and consumer electronics) have comparatively high storage costs, as they are optimized for power rather than energy. This premium is acceptable for mobile applications, but a distributed stationary application for this technology will not likely be cost efficient.

3.4 Pumped hydroelectric storage

Pumped hydroelectric storage is an established, mature technology. Given its mode of operation (in periods of low demand, low-cost electricity is used to pump water to a reservoir at a higher elevation; this water is subsequently released through turbines in peak load periods to generate hydroelectric power), it requires very specific locations: a suitable location must have sufficient water and a significant difference in altitude between its high and low points. Such geographic constraints, combined with political resistance (the environmental impact of the construction of a pumped hydroelectric storage site is significant), pose clear limits to the additional growth potential of the technology.

Yet, some experts regard selected pumped hydroelectric plants, connected by long-distance transmission lines, as potential buffers for fluctuating renewables. A prominent pointer toward such developments is the NorNed cable, which transports up to 700 megawatts between hydroelectric plants in Norway and the Netherlands. There may be a few more new pumped hydroelectric projects in continental Europe – notably in Austria, Norway, Spain, and Switzerland – within the next 10 to 15 years.[9] But local and political resistance associated with the environmental impact of these plants will remain a large barrier.[10]

Given the strengths and weaknesses of these different technologies and their different applications, we believe that a combination will ultimately be used to provide the necessary compensation capacity. None of the technologies listed is capable of solving the challenge alone.

4 New business opportunities will emerge, but so will challenges

Growing demand for storage will present a range of business opportunities for energy companies, utilities, and related players. But it will also require gaining clarity on some key issues. We answer the fundamental questions below.

4.1 Who are the likely operators of storage facilities?

The value chain for energy storage facilities encompasses four distinct roles: *component suppliers,* which will design, produce, and market components for storage facilities; *solution providers,* which will design, produce, and market components as well as design and integrate systems and storage facilities; *storage facility operators,* which will operate facilities as well as buy and sell electricity; and *service operators,* which will offer maintenance and operational assistance to storage facilities. The provision of components, solutions, and service is similar to existing businesses in related applications (such as gas storage, power plants, water technology, and industrial facilities), but the actual operation of energy storage facilities presents a new set of challenges to potential operators and investors.

By necessity, the operators of energy storage facilities will most likely derive from the energy sector, as only these players have applicable experience that can be leveraged.[11] Thus utilities, municipalities, independent power producers (IPPs), and grid operators would all be immediate candidates. The European Union's unbundling directive, though, prevents grid operators from operating storage facilities, because it classifies grid operators in the "generation" category. Hence the field seems to be left to utilities, municipalities, and IPPs.

---

[9] Iberdrola is currently constructing an extension of its La Muela pumped hydroelectric plant (La Muela II will have 850 megawatts of capacity and is expected to be completed in 2012) and plans an additional 900-megawatt facility at Alto Tâmega in Portugal, scheduled for completion in 2018.

[10] The proposed North Sea supergrid appears to be an exception: it would use hydropower from Norway as a complementary energy source for offshore wind parks. In the current planning, however, this relates mainly to the country's existing hydropower stations, which can modulate their generation output but do not have pumps. Thus, we do not consider them storage facilities vis-à-vis those discussed in this paper; rather, we view them as a means of intermodal compensation.

[11] It is possible that households could eventually adopt their own storage solutions. From a cost perspective, this would be the most expensive solution. We do not focus on this possibility in this paper.

Our research suggests that large established utilities may be reluctant to take on the role of storage operator, given their traditional focus on conventional investment projects, which compete for internal funding. However, smaller utilities, especially municipal ones, could turn out to be more flexible and open to exploring this new line of business if they believed it offered them an opportunity to strengthen their local market footprint. Whether it will afford them this opportunity will greatly depend on their specific regional or local setting.

IPPs also have incentive to look seriously into energy storage opportunities: as the share of fluctuating renewables in overall generation continues to rise, regulators are expected to increase IPPs' responsibility to better structure their power feed-in relative to peak and off-peak times. Given their small size, however, many IPPs may struggle to bear the required investment alone. Hence, they may need to form alliances to have the capacity to implement large storage projects, and single IPPs may provide storage for smaller, decentralized projects (such as small communities) if required.

It is also possible that other players will enter the storage market once it has become established. For operators of underground natural gas storage, for example, moving to A-CAES would not be a huge leap. New entrants could also derive from other quarters. Suppliers of PV equipment are in fact already offering small-scale battery storage solutions to improve their products' reliability.

4.2 What are the economic incentives for operators?

The key lever for the economically successful operation of an energy storage facility is the spread between the cost the facility incurs for charging energy and the price the facility can obtain for discharged energy. This price spread has to cover not only capital expenses as well as operations and maintenance costs, but must also make up for efficiency losses incurred during the storage process. Our calculations indicate that, for a while, this spread will not be sufficiently large on a continuous basis for operators. It is also extremely complex for operators to reliably chart their business cases, as the actual utilization of a storage facility at any point in time depends on the complex interaction of various parameters, such as weather, load, and grid capacities. While additional revenues from the provision of reserve energy and other grid services may help the financial viability of the storage business model at some point, the economics of storage will remain unattractive, or at least not fully transparent, for some time.

We should mention that providing grid stability and transmission and distribution deferral is currently a viable business case for storage operators, especially in the US.[12] Since this is a transitory business case that will not require large volumes of storage, we do not analyze it further here.

[12] Transmission and distribution (T&D) deferral refers to the use of storage infrastructure to defer the replacement or upgrading of existing T&D equipment that is no longer capable of handling peaks in the grid: by using storage for peak shaving, existing installations can be used for a longer period of time. The business case for the storage investment is driven by the cost of capital required for the T&D upgrade and by the number of years the upgrade can be deferred.

Given the currently weak business case for storage technologies compared with other approaches to compensation, it is unlikely that investments in storage facilities will be made in the next few years. However, storage is – as concluded – a vital component of the push toward fluctuating renewables. Unless the integration of fluctuating renewables into the grid is provided for, the growth of power generation from intermittent sources (i.e. wind and sun) will reach a ceiling. Many governments have committed to ambitious targets concerning renewable energy, so they will need to provide suitable forms of incentives and subsidies in order to assure market readiness for storage infrastructure and to get key pilot projects off the ground. These could be investment-linked incentives (for example, loan guarantees and investment tax credits, such as those widely used in the US) or power-price-linked incentives (for example, guaranteed feed-in tariffs or price premiums paid per unit of energy stored in proportion to subsidies given to renewable energy, such as those common in many European countries).

In addition to these "pull" incentives, "push" incentives could be employed. Today, feed-in tariffs for renewable generation are time independent. By adding a peak-related component, owners of renewable generation facilities might be pushed to structure their feed-in to optimize their returns, increasing their need for storage. Alternatively, regulators could require new installations of wind or solar PV to provide a certain amount of storage to structure their output.

4.3 What are the key success factors for operators?

To succeed, storage operators will need to make the right decisions about strategic positioning and technology. Concerning strategic positioning, operators should aim to optimize the following:

Degree of utilization. Storage projects should be designed to be available for frequent charging and discharging. This will require a suitable location, sufficient input and output capacity, and a high-performance grid connection.

Revenue streams. The operating model should facilitate revenue creation not only from the structuring of fluctuating renewables' generation, but also from other sources. For example, some part of overall capacity could be set aside to provide reserve capacity as a revenue source.

Use of subsidies. Storage facilities should be designed to leverage existing (and forthcoming, if known) subsidies. This will necessitate a close monitoring of regulatory developments.

Optimization of technology will require focusing on these considerations:

Size and capacity. Operators will need to strike a balance between size of investment and potential revenue streams. A larger incremental investment could, for example, enable the parallel provision of reserve capacity.

Charging and discharging capacity. To enable high turnover and the ability to compensate for day-night fluctuations, input and output capacity must be sufficiently dimensioned to ensure the possibility of charging and discharging within approximately 10 hours.

Response time and flexibility. The provision of reserve energy, for example, requires a response time of less than 15 minutes. The technology must enable this.

Charging energy prices vis-à-vis efficiency losses. Technologies that have low efficiency, for example hydrogen, are better suited for settings that command very low prices for charging energy than for settings in which substantial prices have to be paid for charging energy – as long as self-discharge, or the spontaneous loss of energy, is low (which is the case for hydrogen).

We are witnessing a paradigm change. Historically, electricity generation has been designed to follow demand. Now, we are moving toward a world in which virtually continuous demand is expected to be met, to a large extent, by energy sources – wind and sun – that are not "on," or able to generate, much of the time. This challenge makes electricity storage critical – and the next frontier in energy infrastructure.

While the business case for investing in storage is currently weak, that situation will necessarily change. Today's fluctuations in generation are compensated for relatively easily and cheaply by flexible conventional power plants, but there are limits to how much capacity these plants will be able to provide. Simultaneously, the march toward a fossil fuel-free energy landscape continues: ambitious targets for the share of electricity to be provided by renewables have been formulated and confirmed. Wind and solar PV are the most competitive and widely available renewable sources and will certainly account for the lion's share of renewable energy produced – and they require storage to be viable.

To realize fruit from their investments (through subsidies) in building a renewable generation base, governments will necessarily turn their attention to integration. And compensating for the fluctuations induced by wind and solar power generation, which represent a big piece of the integration challenge, will require massive investments. A tremendously large market will develop around fluctuation management technologies, and large-scale electricity storage will account for a good part of those investments. Indeed, we expect a strongly growing market around electricity storage, with annual revenues well above 10 billion EUR by 2020. Already, the beginnings of this market are evident: the first projects are being tendered, research activities are picking up, and venture capital is viewing electricity storage as a key investment topic in clean technology.[13] All participants in the industry value chain would therefore be well advised to take the initiative and start positioning themselves for the inevitable. The race for the best technologies, the most advantageous sites, gaining operational experiences, and shaping standards and policies is open, and the commercial payoff for getting into this race early stands to be significant.

Governments that are committed to fluctuating renewables need to be similarly proactive. Given its current economics, storage needs a push to get it off the ground, and governments are uniquely positioned to provide that push. As was the case with renewable electricity *generation* some years ago, the right and timely incentives

[13] Thomson Reuters survey of 41 clean-technology venture capital investors, presented at the World Economic Forum in Davos, Switzerland, in January 2010.

will spur private investment. Such investment, in turn, will ensure that adequate storage capacity exists to allow for the realization of gains from those earlier heavy investments in renewable capacity – and that the vision of a truly sustainable electricity supply can become a reality.

5 The Transition to Greater Penetration of Fluctuating Renewables—and the Near-Term Impact on Energy Generation

The road to an energy landscape dominated by fluctuating renewables will not be without its challenges. The biggest challenge will be to compensate for the fluctuations that characterize electricity generation from these technologies. It should also be noted that as wind and solar photovoltaic technologies continue to evolve – wind being increasingly deployed offshore and centralized, solar PV being increasingly deployed on rooftops and distributed – they will pose very different challenges in terms of compensation, with significant implications for the underlying transmission and distribution grids.

Four types of fluctuations can be distinguished. Day-night fluctuations are specific to solar PV, while both wind and solar PV display distinct seasonal and annual patterns, depending on region. For instance, in Europe, there is more wind in winter and more sun in summer, so overall fluctuations are somewhat lower than elsewhere. Both types of fluctuations are cyclical and predictable, and they even coincide with load patterns to some extent. But fluctuations due to short-term and medium-term changes in weather (stormy vs. calm, sunny vs. cloudy) are much more erratic and therefore difficult to forecast. This is why the so-called forecast error continues to pose a challenge for generation dispatchers.

As a result of these fluctuations and the compensation challenges they pose, renewables' growth will have severe impact on existing generation systems. In countries with constant or even declining energy demand (for example, most industrialized nations), as capacity for fluctuating renewables grows, the same amount of energy will be generated by a significantly larger fleet of power plants.

Since power from renewable sources will receive feed-in priority, the utilization of conventional power plants will decrease.[14] Hence the residual capacity that can be dispatched across the remainder of the fleet will be low, especially in high-wind, low-load periods. It will always be necessary to operate a number of must-run conventional plants to ensure a supply of reserve capacity so that the remaining conventional plants can operate only very intermittently. But it is these must-run power plants that need to be able to provide flexibility to buffer variations in output from fluctuating renewables.

[14] This is so except in currently rare extreme cases when grid stability is threatened and feed-in from renewable sources can be stifled.

This will have a threefold negative effect on existing power generation systems:

Reduced utilization rates will translate into a higher levelized cost of energy (LCOE), owing to the greater weight of fixed costs per generated megawatt hour. This will become more acute as the share of fluctuating renewables continues to rise.

To provide the flexibility required to balance the feed-in of fluctuating renewables, generation systems will have to shift from base load plants, such as nuclear and lignite plants, that have low variable costs, to more flexible mid load and peak load plants, such as combined-cycle gas turbine (CCGT) and gas turbine (GT) plants, which have higher variable costs.[15]

Adding "priority feed-in" capacity in the form of wind and solar PV will likely have a negative effect on wholesale power prices and thus on generation margins. This will have a systemic effect, as the "cheaper" power plant will become the marginal one, setting the price for all others. Energy price volatility will also increase substantially.

While it is therefore possible that the generation landscape will evolve to a new optimum over time as the penetration of fluctuating renewables grows, during the transition period, excess costs will be incurred, and returns on currently invested capital will be diminished.

[15] To some extent, modern hard-coal power plants can also be operated as rather flexible mid load units; however, their efficiency rate is lower, and carbon dioxide emissions are higher than those of plants powered by CCGT and GT.

Balanced growth through local entrepreneurship: The Komba coffee project in southern Tanzania

Felix Meissner and Xaver Kazimoto Komba

Abstract:

This case study describes the approach and learnings of the *Komba Coffee Project* in Songea, Tanzania. The social start-up, which is driven by local entrepreneurs from the NGO *Songea Network Center*, aims to improve coffee farmers' economic conditions by allowing them a larger share of the value-creation process and giving them bargaining power. Thanks to the re-designed supply chain and the direct export to roasters, farmers have a greater share of the value added and, thus, the profits. In line with Banerjee and Duflo (2011), this case study argues that regional socioeconomic deficits can be successfully addressed when their specific characteristics are taken into consideration. For sustainable growth, economic activities should be supported by the important stakeholder groups and driven by local citizens: Balanced growth is ultimately a question of creating self-confident communities with project ownership and leaders committed to sustainability. Finally, developing a project that is financially self-sufficient and sustainable takes time: Citizens should not only understand the concept and its intentions, but also accept responsibility and ownership for it to be a success.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*, Management for Professionals, DOI 10.1007/978-3-642-24653-1_12,

Table of Contents

1 Africa: A continent full of opportunities and potential

According to the McKinsey Global Institute (2010), African economies are on the move, especially industries like construction, telecom, banking, manufacturing, mining, and retail have flourished since the 1990s after two decades of almost stagnation. Africa is currently among the fastest growing regions in the world and has about 60% of the world's total uncultivated, arable land. The continent has a gross domestic product that roughly equals Brazil's or Russia's. Africa's rapid development is emphasized by its current future growth expectations: In 2020, Africa's collective GDP will rise to 2.6 trillion USD and consumer spending will be 1.4 trillion USD. There will be 1.1 billion Africans of working age in 2040 and 50% of the population will be living in cities by 2030 (McKinsey Global Institute, 2010). However, despite these bright growth perspectives, most African countries face serious challenges, including poverty, diseases, poor water quality, and insufficient food supply.

2 Tanzania: A country on the rise

Tanzania will play a key role in the future African development and will serve as an important economic and transportation hub in East Africa. With a total population of about 44 million and a gross national income per capita of 1,344 USD (in 2010) (National Bureau of Statistics, 2011; United Nations Development Programme, 2010), Tanzania no longer ranks among the least developed countries. It has slightly but continuously improved its position in the UN's Human Development Index over the last decades (United Nations Development Programme, 2010).

Tanzania has rich natural resources, including gold, natural gas, coal, iron, uranium, platinum, as well as fishery, arable land and dozens of National Parks like the Serengeti, the Ngorongoro Crater, and Mount Kilimandjaro. Moreover, Tanzania is a country that is politically rather stable. Several reforms and improvements in the education system, new import and export regulations, taxation reforms, the creation of investment opportunities, incentives and further policies to attract and protect foreign investments will further foster economic growth (Kulindwa, 2002).

This trend is nourished by the rising global demand for resources, increased international investments, a better-qualified labor force, and increasing international tourism. Tanzania has a great deal of growth potential besides exploiting its natural resources. Apart from mining and agriculture, the service and manufacturing sectors are expected to grow, strengthened by political stability and foreign direct investments.

Tanzania is positioned strategically in the center of the East African coast. With its three main harbors in Dar es Salaam, Mtwara, and Tanga it not only connects its economy to the world, but also that of its neighboring countries Mozambique, Malawi, Zambia, the Democratic Republic of Congo, Burundi, Rwanda, Uganda, and Kenya. The Tanzania Ports Authority has announced plans to further expand their capacity, especially in Mtwara (Tanzania Ports Authority, 2011).

But will the expected growth and development really be for the benefit of Tanzanian citizens? Or will only a few actors benefit from the country's natural resources, exports, and development as was the case in other resource-rich African nations?[1] Will the expected boom on the African continent really be for the benefit of all its people and not only for a few political and economic chiefs?

If one examines economic research, there are many recommendations on how to use Africa's untapped potential to secure developments that will improve the economic situation on a broad level. However, as Banerjee and Duflo (2011) conclude in their recently published book "Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty", there is no master recipe or magic formula that fits all nations, all their economies, and all their problems. In line with our personal experience, it is worthwhile analyzing and thinking about this particular problem since specific answers are needed in the majority of cases.

In line with this train of thought, the following case study does not intend to provide a one-size-fits-all solution, but rather describes the activities and insights of the authors and their team, who have been actively involved in numerous development initiatives in southern Tanzania. In order to foster sustainable and balanced growth, the inclusion of local needs and the co-development of initiatives together with the local citizens are pivotal. Only when people can be convinced to and are willing to advance development themselves, can such growth initiatives enjoy long-term success.

The rest of the chapter is structured as follows: The second section provides a brief overview of the Tanzanian NGO Songea Network Center's philosophy and activities. It illustrates that the vision and mission drafted by its founding members take a holistic perspective on improving the regional conditions by operating in a network with the local citizens. The third section looks at the entrepreneurial Komba Coffee project and shows that this locally initiated initiative is in many ways a paramount example of a co-developed model of sustainable regional development that takes various stakeholder needs into consideration. The fourth section concludes by highlighting several lesson learned that could also be important for other projects that wish to pursue sustainable and balanced development.

3 SNC: Fostering socio-economic development

Songea Network Center (SNC) is a Tanzanian NGO founded in 2005 by several local entrepreneurs and headed by Xaver Kazimoto Komba in Songea, Southern Tanzania. The aim was to institutionalize a platform of and for social entrepreneurs to support their idea generation and implementation of projects to help the local population. SNC is supported by many social investors, both from Tanzania and Europe. However, a major principle is to run projects that are based on economic principles and to make activities less dependent on donations. SNC is well con-

[1] See, e.g., Kronenberg (2004); Sachs and Warner (2001).

nected to political, social, religious, and business actors in the Ruvuma district. According to the organization's vision and mission statement, there are five key pillars on which activities focus:[2]

3.1 Entrepreneurship

Regardless of whether farming programs, school projects, or cultural tourism programs are initiated, all activities are strongly shaped by SNC's entrepreneurial spirit and thinking. As a network of entrepreneurial personalities, SNC believes in the power of entrepreneurship. SNC has often taken the entrepreneurial risk of being the first mover to find solutions for local problems. Before their kick-off, many projects addressing social or economic deficiencies seemed unlikely to succeed, but proved to be effective and were then multiplied with cooperating partners or imitated by others.

The Komba coffee project markets fairly traded coffee for the local coffee farmer cooperatives of the Songea-Mbinga region in Europe. This activity, as well as several other agricultural projects and the first Internet cafe in southern Tanzania comprise the SNC entrepreneurship pillar.

3.2 Health

Health issues, especially in the Songea rural surroundings, are a considerable problem. Villages are often far from the nearest hospital ward. Medicine and (public) transportation possibilities are scarce and expensive. Health problems include malaria – which is usually worst during the rain seasons when the road conditions are at their worst – and infections, often caused by stomach problems, tuberculosis, and HIV/Aids. In order to mitigate these problems, SNC built a hospital ward in the Mbinga-Mharule village by means of a public-private partnership, which sponsored a nurse and a midwife, and organized health education programs. There are also initiatives to improve water quality and supply by constructing and maintaining sunk wells.

3.3 Education

SNC's entrepreneurs believe that investing in education is key for the region's long-term development. Since the founding of SNC in 2005, two kindergartens and a secondary school for about 400 students have been successfully started. Further, Songea Network Center is Southern Tanzania's pioneer regarding new technologies: Even before SNC's official founding, the Valongo Internet Cafe was established in 2002 as a service to the local population. For some years, the internet cafe was not only the first, but also the only public internet access in southern Tanzania. In response to the strong general interest shown in the World Wide Web, computer and internet training programs were offered.

[2] Owing to the fast pace of development, SNC works with multiple leaders who deal with the issues emerging from project teams.

3.4 Cultural Exchange

SNC believes that, in a globalized world, mutual understanding and respect have become increasingly important for successful partnerships. Cultural exchange does not only mean inviting European partners and sponsors to Tanzania, but also sending out own team members to allow them to experience and learn from cultures outside Africa. Over the last few years, SNC has welcomed and hosted several groups concerned with ecological and cultural activities.

3.5 Environment

Environmental protection is becoming an increasingly important topic in sub-Saharan Africa, especially since the World Bank and other institutions predict that climate change will negatively affect agriculture in the future. SNC offers training programs and discusses cultivation, water management, and climate change issues with farmers. SNC has also supervised a tree-planting project to counter desertification. Overall, SNC is committed to thinking globally and taking the needs of various stakeholders into consideration, while simultaneously acting locally by addressing very specific socioeconomic issues in Ruvuma.

4 Development through local entrepreneurship: The Komba coffee project

Initiated by Florence Komba, Thadei Chiwangu, Michael Mbano, Xaver Komba, and Felix Meissner, the Komba coffee project – Komba is a frequently found and typical Tanzanian name – is an important flagship project that brings positive cash inflows and serves as a platform to communicate the Songea Network Center's social initiatives in Europe.

4.1 Origins and ideation

When SNC started its coffee project, there were already several foreign-driven fair trade projects across the country. Organizations certified with recognized fair trade labels marketed these high-quality coffee beans around the world. Was another fair coffee initiative supervised by SNC really necessary?

After an analysis of the existing coffee initiatives, it became clear that the reality differed from the promoted image of "fair trade" coffee. At least some of the projects marketing coffees certified as fair and sold in the northern hemisphere could neither track the origins of their coffee, nor name the exact areas, villages or farmers in their program. According to insiders, agencies were often paid for their fair trade certificate without the coffee being purchased fairly and transparently from the producers. Sometimes, no proper assessment was available of the certification processes. Consequently, these business structures neither allowed the producers fair participation, nor did they provide a prospect of advancing the coffee regions in the long run. These projects also had only a few beneficiaries.

Farmers and their cooperatives approached SNC, as they felt that their chances of selling their products at good prices and exporting them were rather limited. Market inefficiencies and regional monopolies – since only certain brokers had a permission to purchase and export coffee beans –, as well as a lack of proper legal protection were major problems for the farmers. They actually had very limited market access via the merchants purchasing their coffee and reselling it at the national coffee auctions in Moshi. A merchant usually buys unprocessed coffee beans rather cheaply from farmers, processes these beans centrally with his own or rented machinery, and resells the beans – either at the national coffee auctions or directly to customers abroad. The margins are quite attractive for such coffee brokers since their processing, transportation, and storage activities are further value-adding steps.

However, owing to this system, the farmers had little room for negotiation, especially since they had few other options for selling their coffee. Moreover, regardless of the system, they always faced risk and insecurity since their sales were never guaranteed. Would the merchants again purchase their coffee the following year? Would coffee prices drop? Since coffee is their only product, thus securing their income, the inefficient markets and the brokers' strong bargaining position, exacerbated their problems. The farmers' bargaining position was so poor that they often had to accept the merchants' small prepayments, thereby financing the coffee beans' resale at an increased personal risk. Large wholesale merchants located many hundreds of kilometers away had at times only picked up the coffee for a small prepayment, but never returned to pay the outstanding amount. These deceived farmers had little legal security. Without the means to pay lawyers and with only a very basic education, they had no opportunity to enforce their rights. Another upcoming risk is the ownership of land. Although farmers traditionally own their coffee fields, their land was never officially (= legally) registered at the land ministry. In Northern Tanzania, an international coffee company registered the fields that certain coffee farmed had owned for many generations and expelled them.

4.2 Starting the Komba coffee project

When the managing board of SNC was asked to mitigate the farmers' situation, it became clear that an entire re-design of the production and marketing system was required for more value creation and bargaining power. The farmers were usually already loosely organized in village communities and these existing local structures were used as a basis to organize the farmers.

The re-designed value chain, which the partnering coffee-growing villages and the SNC own, works as follows:

- Even before the harvesting season starts, SNC is present in the villages, discussing the farmers' situation and ambitions with them. Agreements to buy that guarantee minimum prices are signed in advance.
- In order to signal the SNC's strong commitment, some producers even receive payments in advance to mitigate private financing issues.

- Building further on the SNC's strong presence and positive image in the district – thanks to past successful agricultural projects –, a permanent office has been opened close to the coffee villages to give everyone the opportunity to come and ask questions. This permanent presence gives SNC increased credibility and provides the farmers with legal security.
- The farmers' participation in SNC's coffee project remains voluntary. However, almost all farmers in the partner villages participate and are represented by a delegate from their specific village.
- In the harvesting season, which starts in March, SNC tries to leave as much value creation decentrally in the villages as possible:
  1. *Picking*. The cherries are individually picked by hand at the peak of ripeness.
  2. *Processing*. SNC has acquired a central processing unit that removes the fruit covering form the beans in a wet process. After fermentation, the beans are further sorted, cleaned, and finally spread out in the sun for drying.
  3. *Milling and storing*. The final steps include the removal of the last layers of the dry skin and a final cleaning, which have to be done centrally by using a bigger rented machine. Finally, the coffee is sorted and stored in bags until it is shipped.
  4. *Shipping and selling*. SNC was granted a license that allows the organization to officially purchase the coffee and export it after a quality check by the Tanzanian Coffee Board. This allows the social entrepreneurs to control the land and sea transportation until the coffee reaches Hamburg.
- SNC founded its own distribution agency to take care of the marketing and distribution in Europe. Owing to a strong and exclusive network of partnering roasters with a strong commitment to the Komba coffee project, SNC could negotiate long-term-orientated sales agreements that minimize market risks. The partner roasters are so committed that they partly finance the harvest and the farmers' remuneration. By circumventing intermediaries and motivated to build a long-lasting Tanzanian-driven business with its partners, SNC can ensure higher reimbursement of the farmers and further investments in the villages.

4.3 Outlook

The success of the Komba coffee project has exceeded all expectations. The few, carefully selected partner roasters' demand for coffee is a multiple of the current capacity. The end consumers do not only cherish the coffee for its fruity flavor, but also for the farmers and SNC's entrepreneurial spirit and transparent value chain. Already in its fourth season, the Komba coffee project could continue to grow and ensure sustainable development with improved conditions for the farmers.

Although already successful, SNC's entrepreneurs and their high ambitions there still have a long way to go. In order to make the coffee business a driver of bal-

anced growth in the region, continuous improvements will be also required in the future. Current plans seek further advances by

- offering the farmers training and education to increase their knowledge of farming, especially of organic production and optimal harvesting.
- acting as an advocate to officially register farmers' lands given the increasing number of land-grabbing incidents.
- further developing the rural region's untapped opportunities by, for example, providing basic health insurances and offering farming families better access to education.

5 Lessons learned: Thinking globally, acting locally

This case study has outlined that regional socioeconomic deficits can be successfully addressed if the specific problem characteristics are taken into consideration. The Komba coffee project is still on a rather small-scale, basic level. Our (still preliminary) learnings include the following points:

5.1 People driven

Activities that should be sustainable in the long term must be wanted and driven by the local citizens. If initiatives are developed and owned by the local people, they are more likely to be sustained. Problems can be addressed more specifically and ownership boosts motivation (Kulindwa, 2002). In order to achieve this, we need more African entrepreneurs committed to reducing poverty, more public-private partnerships, and, overall, more support from state institutions.[3]

5.2 Financially self-sufficient partnerships

Africa has for too long been viewed as a charity case instead of a dynamic continent that can soon become an important economic engine (U.S. Department of State, 2011). Africa is a rich continent. It should not be treated as a continent incapable of doing business with the rest of the world, but rather as a worthy partner. Balanced growth is ultimately a question of creating self-confident nations whose leaders are committed to sustainability. Many development initiatives are not only driven by foreign or state institutions, but are also financed by them. This often counteracts long-term success. Mere external sponsorship makes developing countries revert to a relationship of dependency instead of allowing them to be regarded as equal partners. A key concern of SNC's entrepreneurs is not only to run projects with a deep social impact, but which are also (financially) independent from donators and state institutions.

[3] See also U.S. Department of State (2011); Dalohoun et al. (2009); Mc Lennan and Ngoma (2004); Thomas-Slayter (1992); Valente (2010).

5.3 Convince all stakeholder groups

Even well-financed initiatives supported by the local citizens can fail because other, important groups have not been included and convinced. SNC pays close attention to intense communication with the state and state-related supervising authorities. The Tanzania Coffee Board (TCB), which is responsible for quality controls and for export permission, acts as a facilitator and promoter of the Komba coffee initiatives. Since state institutions are expected to have more influence in the future, it is very wise to cooperate with them.

5.4 Growing slowly

Haraka haraka, haina baraka: Growing in a sustainable way means growing slowly. Haraka haraka, haina baraka is Swahili for "hurry does not bring any blessings." SNC's agricultural projects have shown that rushing does not bring long-term success. If projects are run by local farmers, they have to be absolutely convinced about the project, its processes, and benefits even before starting. It took SNC more than four years to communicate its intentions and to build trust with the coffee villages. Only now, after three seasons of successful coffee exports, are the farmers beginning to trust SNC's good intentions and competencies, and are a greater number of villages and farmers willing to cooperate. Further, the sound redesign of the value chain also required time.

5.5 Strengthen local ownership

The Komba coffee project initiative is one of the first fair coffee initiatives initiated and run by local Tanzanians. Many fair projects were previously and currently started by foreign NGOs and were then controlled by expatriates who pursued their own objectives. Projects with a de facto local ownership have a completely different locally anchored legitimization, and their governance mechanisms depend less on foreign control.

References

Banerjee, A. V. and E. Duflo (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. Toronto: Perseus.

Dalohoun, D. N., A. Hall, and P. Van Mele (2009). Entrepreneurship as driver of a self-organizing system of innovation: The case of Nerica in Benin. *International Journal of Technology Management & Sustainable Development* 8(2), 87–101.

Kronenberg, T. (2004). The curse of natural resources in the transition economies. *Economics of Transition* 12(3), 399–426.

Kulindwa, K. (2002). Economic reforms and the prospect for sustainable development in Tanzania. *Development Southern Africa* 19(3), 389–403.

Mc Lennan, A. and W. Y. Ngoma (2004). Quality governance for sustainable development? *Progress in Development Studies* 4(4), 279–293.

McKinsey Global Institute (2010). Lions on the move: The progress and potential of African economies. Report, McKinsey Global Institute.

National Bureau of Statistics (2011). National Bureau of Statistics Tanzania – NBS, Project Website.

Sachs, J. D. and A. M. Warner (2001). The curse of natural resources. *European Economic Review* 45(4-6), 827–838.

Tanzania Ports Authority (2011). Tanzania Ports Authority, Project Website.

Thomas-Slayter, B. P. (1992). Implementing effective local management of natural resources: New roles for NGOs in Africa. *Human Organization* 51(2), 136–143.

United Nations Development Programme (2010). Human development index and its components.

U.S. Department of State (2011). Remarks at the corporate council on Africa's Seventh Biennial U.S.-Africa business summit.

Valente, M. (2010). Public and private partnerships for sustainable development in Africa: A process framework. *Journal of African Business* 11(1), 49–69.

Editors

Giulia Mennillo is a Research and Teaching Assistant at the Institute of Economics at the University of St.Gallen, Switzerland. She is a doctoral candidate in International Affairs and Political Economy. Her main research interests deal with the economic weight of states and its consequences for the distribution of relative political power in the international system. Giulia Mennillo studied International Economics at the Eberhard Karls Universität Tübingen and at Bocconi University in Milan, Italy. She coordinated part A, "Balanced growth – the economics perspective," in this book. She is responsible for the corresponding panels at the 9th Docnet Management Symposium.

Thomas Schlenzig is a Research Associate and doctoral candidate at the Strategic Management Chair of the Institute of Management at the University of St.Gallen, Switzerland. His main fields of research are organization strategy and structure. Thomas Schlenzig studied Business Administration at the University of Mannheim in Germany and at Linköping University in Sweden. He coordinated part B, "Balanced growth – a business perspective," in this book. He is responsible for the corresponding panels at the 9th Docnet Management Symposium.

Elmar Friedrich is a Research Associate and Ph.D. candidate at the Good Energies chair for Management of Renewable Energies at the University of St.Gallen, Switzerland. His dissertation research focuses on analyzing strategic investment decisions regarding renewable energies of utility companies. Elmar Friedrich was a member of the Swiss delegation at the COP15 UNFCCC Climate Negotiation in Copenhagen in 2009 and teaches courses that link climate change and business. His dissertation project is supported by a Nagelschneider Foundation scholarship. He coordinated part C, "Balanced growth – a resource management perspective," in this book. He is responsible for the corresponding panels at the 9th Docnet Management Symposium.

G. Mennillo et al. (eds.), *Balanced Growth: Finding Strategies for Sustainable Development*,
Management for Professionals, DOI 10.1007/978-3-642-24653-1,

List of Authors

Hans Christoph Binswanger is a Swiss economist. His research field has focused on monetary theory, and environmental and resource economics. He has developed an environmental tax reform concept. Hans Christoph Binswanger studied economics in Zurich and Kiel. He gained his doctoral degree from the University of Zurich and was awarded his habilitation by the University of St.Gallen. He lectured as a tenured professor of economics at this university from 1969 to 1994, when he was awarded emeritus status. From 1967 to 1992, he was the director of the Institute of Economics (FGN-HSG). From 1992 to 1995 he served as director of the newly formed Institute for Economy and the Environment (IWÖ-HSG).

Thomas Davoine is a Research Assistant and Ph.D. candidate in economics at the University of St.Gallen. He studied applied mathematics at the École Polytechnique Fédérale de Lausanne and at Rutgers University. He has, for several years, also worked for management consulting firms, advising clients in the international public sector, humanitarian aid agencies, and the banking industry.

David J. G. Dwertmann is a Research Associate and doctoral candidate at the Center for Disability and Integration (CDI-HSG) at the University of St.Gallen, Switzerland. He studied psychology with business administration as a minor at the universities of Mannheim (Germany) and Bern (Switzerland). Before joining the Center for Disability and Integration, he conducted research at the University of Mannheim and San Diego State University. His research interests focus on diversity, organizational culture and climate, organizational identity, and program evaluation.

Daniel Kienzler is a Research Assistant and Ph.D. candidate at the University of St.Gallen. He has attended Ph.D. courses at the Swiss National Bank's Study Center Gerzensee. Daniel Kienzler studied economics at the University of Hohenheim in Germany and at the McMaster University in Hamilton, Ontario with an emphasis on econometrics and monetary economics. His research interests focus on macroeconomics and monetary policy.

Xaver Kazimoto Komba studied Philosophy at the Peramiho Major Seminary in Songea, Tanzania and Theology at the Royal University of Santiago de Compostela, Spain. As a serial entrepreneur and priest, he gained extensive experience in managing social and economic projects. He founded the Parish Mjimwema in the Archdiocese of Songea, the Sapientia Secondary School in Songea, and initiated several agricultural projects and education initiatives. He is founder of the Songea

Bavaria Netzwerk e.V. in Germany, the NGO Songea Network Center Ltd. in Tanzania, and Komba Coffee Tanzania GmbH in Germany. Xaver Komba is currently a Ph.D. student at the Ludwig-Maximilians-University in Munich, where he is writing his thesis on "The sustainability of the African church: A sociological and theological analysis with reference to small Christian communities in Tanzania." His expertise and main research interests lie in the areas of climate change, social entrepreneurship, project management, and the sustainability of small communities in Tanzania.

Markus Kreutzer is an Assistant Professor of Strategic Management at the University of St.Gallen, Switzerland, where he also serves as Executive Director of the Excellence Initiative on "Responsible Corporate Competitiveness" (RoCC), a multidisciplinary initiative. His research interests are in the area of strategy process, strategic renewal, organizational control and coordination, and alliance strategy.

Justus Julius Kunz is a Research Associate and doctoral candidate at the Institute for Leadership and Human Resource Management (I.FPM) at the University of St.Gallen, Switzerland. In addition, he also works as a consultant at the energy factory St.Gallen AG. He studied Psychology and Communication Science at the University of Erfurt and was part of different interdisciplinary research groups (e.g., at the Max-Plank-Institut, Bonn) investigating the social psychological aspects of decision-making processes. For his master thesis he evaluated a change project for a leading public utility company. His research interests focus on strategies to avoid company-wide over-acceleration.

Felix Meissner worked in Tanzania for the Missionary Benedictines in Peramiho and Ndanda, as well as on several development projects for the Archdiocese of Songea. Thereafter, he received a Bachelor of Arts in Corporate Management & Economics with a minor in Communication & Cultural Studies at the Zeppelin University in Friedrichshafen, and a Master of Arts in Accounting & Finance at the University of St.Gallen. During his studies, he worked for several start-ups, corporations, and research institutes in the fields of accounting, business development, and strategy consulting in China, Germany, Sweden, Switzerland, and Tanzania. He co-founded Songea Bavaria Netzwerk e.V. in Germany, the NGO Songea Network Center Ltd. in Tanzania, and Komba Coffee Tanzania GmbH in Germany. Felix Meissner is currently a master and Ph.D. student at the University of St.Gallen. His main research interests lie in the areas of alliance and network strategy, the organizational design of cooperations, and social entrepreneurship.

Liudmila Nazarkina is an oikos Ph.D. Fellow in Strategy and Management at the University of St.Gallen. Her research is focused on sustainability-driven firms and their growth strategies. She holds a Master's degree in Strategic Management from the State University – Higher School of Economics in Moscow and a Master's degree in Corporate Social Responsibility from the Université Paris XII. Besides her academic activities, Liudmila has worked in sustainability consulting in France

and the UK. She currently leads oikos projects related to teaching cases in sustainability, management, and entrepreneurship.

Melanie Oschlies is a consultant at Bain & Company in Zurich. Currently, she is pursuing her Ph.D. in the area of renewable energy management at the University of St.Gallen. Before joining Bain, she studied business administration at the University of St.Gallen, as well as at the HEC Montréal and the National University of Singapore.

Cornelius Pieper is a principal in the The Boston Consulting Group's (BCG) Cologne office. He is the BCG topic expert on energy storage and has worked extensively on this topic. Before joining BCG, Cornelius Pieper worked as a researcher at the Geographical Institute of Bonn University where he completed his Ph.D. (Dr. rer.nat.).

Friedemann Roy is a Senior Housing Finance Specialist at the World Bank. His work is focused on an appropriate enabling environment for housing finance, as well as adequate and corresponding finance and investment mechanisms in emerging markets. With proven business development skills in investments and project management skills in emerging markets for more than 10 years, in particular in central and Eastern Europe, Asia and Africa, as well as BRIC countries, he has contributed to investment recommendations, including strategy, design, as well as development and implementation of appropriate investment channels (especially funding long-term assets). Previously, he held positions at the International Finance Corporation (IFC) in South Africa, the Frankfurt School of Finance & Management in Germany, and at Commerzbank AG in Frankfurt, Paris, London, and Warsaw. In addition, he worked as an editor for the "Housing Finance International" Journal, which is published by the International Union for Housing Finance (IUHF).

Holger Rubel joined the Frankfurt office of The Boston Consulting Group (BCG) in 1997. He is the Global Head of the sector Green Energy Technologies and a member of BCG's internal Practice Areas Technology & Communication and Energy. In addition to his work in renewable energies (wind, solar, and energy storage), Holger Rubel primarily supports clients in industrial high-tech sectors such as automation technology, industrial IT systems, and infrastructure solutions. Holger Rubel studied physics and mathematics at Heidelberg University and Cambridge. He has a M.Phil. degree from Cambridge and a doctorate in semiconductor technology from the Max Planck Institute in Stuttgart.

Michèle F. Rüdisser is an Assistant Professor of Organizational Control and Governance at the University of St.Gallen. She is also a Visiting Professor at the Università Commerciale Luigi Bocconi's School of Management in Milano, Italy. Previously, she was an Assistant Professor at the University of Zürich and a Visiting Scholar at the Department of Business Strategy and Policy at Tsinghua University's School of Economics and Management in Beijing, China.

Jens Schädler is a partner at Bain & Company in Zurich. He has 18 years of consulting experience with a focus on the telecommunication and industrial goods and services sectors. He serves clients on a variety of topics, including growth and portfolio strategies as well as performance improvement. He holds a Master's degree in business administration and international management from the University of Cologne and the Community of European Management Schools (CEMS), and a Ph.D. in law from the University Hamburg.

Uwe Sunde is Professor for Macroeconomics at the University of St.Gallen and director of the Swiss Institute of Empirical Economic Research at the University of St.Gallen. He is also affiliated with IZA (Bonn), CEPR (London), CESifo (Munich), and DIW (Berlin). Uwe Sunde's research is devoted to the theoretical and empirical investigation of the determinants of long-term growth, focusing specifically on the role of demographics and institutions.

Thomas Vischer holds a degree in economics from the University of Bonn and a degree in finance from the University of St.Andrews. He is currently a Ph.D. student at the University of St.Gallen. His dissertation explores the empirical determinants of economic growth, focusing specifically on the role of human capital acquisition.