GERMAN COUNCIL
of Economic Experts

Conseil d'Analyse économique

Monitoring economic performance, quality of life and sustainability

Joint Report as requested by the
Franco-German Ministerial Council

December 2010

GERMAN COUNCIL
of Economic Experts

Conseil d'Analyse économique

Monitoring economic performance, quality of life and sustainability

Joint Report as requested by the
Franco-German Ministerial Council

December 2010

Conseil d'Analyse Économique
113 rue de Grenelle
75007 Paris
Tel.: 0033 1 / 4275 5300
Fax: 0033 1 / 4275 5127
E-Mail: pierre.joly@pm.gouv.fr
Internet: http://www.cae.gouv.fr

Sachverständigenrat zur Begutachtung
der gesamtwirtschaftlichen Entwicklung
Statistisches Bundesamt
65180 Wiesbaden
Tel.: 0049 611 / 75 2390 / 3640 / 4694
Fax: 0049 611 / 75 2538
E-Mail: srw@destatis.de
Internet: http://www.sachverstaendigenrat.org

Released in January 2011
Price: € 15,- [D]
Order number: 7700009-11900-1
ISBN: 978-3-8246-0942-0

Printed by: Bonifatius GmbH Druck-Buch-Verlag, D-33042 Paderborn

PREFACE

1. The Franco-German Ministerial Council decided on February 4, 2010 to ask the French Conseil d'Analyse Économique (CAE) and the German Council of Economic Experts (GCEE) to follow-up on the outcome of the "Commission on the Measurement of Economic Performance and Social Progress" (Stiglitz-Sen-Fitoussi Commission, or SSFC).

The CAE and GCEE have fulfilled this request by preparing a report on

„Monitoring economic performance, quality of life and sustainability".

It discusses how comprehensiveness and accuracy of an indicator set might be traded off optimally with parsimony and cost to provide a reliable basis for regular, timely and digestible reporting on three key issues regarding economic performance, quality of life and sustainability.

2. As the world is emerging from its worst economic crisis of the last six decades, there is a broad consensus among policy makers and the general public that this should be a moment of pause and sincere reflection. From the vantage point of economics and statistics, three intimately related key questions should form the focus of such considerations: First, how can we improve our monitoring of economic performance in order to allow policy makers to gauge the current state of affairs and to react timely and appropriately when crises emerge? Second, how can we broaden our perspective from its current focus on economic performance to an assessment of the quality of life more generally, in order to appreciate what really counts for human welfare? And third, how can we design warning signals that alert us whenever the current manner of organizing our lives endangers sustainability, in order to correct our course of action for the sake of our own future and that of generations to come?

The first and arguably most important conclusion of our study is that a single-indicator approach to measuring human progress is inherently insufficient. Complexity of life and the demands on statistical reporting are too diverse to allow a meaningful condensation of the current state of affairs into a single comprehensive indicator. Instead, we suggest that comprehensive statistical reporting should entail a dashboard of indicators. The dashboard we propose is meant to be a starting point for discussion. It is intended to be rich enough to facilitate a sensible discussion of the relevant facets of human welfare, but it is also not overwhelmingly extensive. Moreover, it provides a balanced representation of the three areas addressed by the key questions, economic performance, quality of life and sustainability. This approach acknowledges that monitoring material well-being is an indispensable prerequisite for sensible economic policy, that life is about more than material well-being, but that human progress in non-material aspects is quite difficult to capture, and that it is wise to take a long-term perspective by outlining the consequences of unmodified human behavior.

3. The two involved institutions prepared this report with the following division of labour: The CAE took the lead in preparing Chapter II and section 2 of Chapter IV, while the GCEE took the lead in drafting Chapter III and section 3 of Chapter IV. Sections 1, 4 and 5 of Chapter IV are a joint product. Chapter I constitutes an introduction and summary of the report.

4. The CAE would like to thank Professor Christian Saint-Etienne for having kindly agreed to be the coordinator for the French Council.

The CAE is also grateful to Philippe Cunéo and Claire Plateau from INSEE for their comments and contributions to this report. The whole staff of the Conseil d'Analyse Économique has helped by providing research and logistic support and must be thanked, especially Christine Carl for editing the French version and Agnès Mouze for documentation.

French contributions owe a lot to the work of CAE's scientific advisers, Associate Professor Jézabel Couppey-Soubeyran, Professor Jerôme Glachant, Professor Lionel Ragot, Professor Stephane Saussier, Professor Thomas Weitzenblum and Associate Professor Anne Yvrande-Billon. They must be thanked for it.

The General-Secretary Pierre Joly can be praised for his contributions and for coordinating this joint report on the French side.

5. The GCEE would like to express his profound gratitude to Professor Dr. Christoph M. Schmidt. His intense efforts as the main author and coordinator on the German side helped immensely in producing the report.

The GCEE would also like to thank staff from the German Statistical Office, specifically from the national and environmental accounts units, for providing helpful comments. As usual the members of the branch that work with the GCEE on a daily basis have helped prepare this report. We would like to thank Anita Demir, Diplom-Volkswirt Wolfgang Glöckler, Diplom-Volkswirtin Birgit Hein, Christoph Hesse, Klaus-Peter Klein, Uwe Krüger, Sabrina Mäncher, Volker Schmitt and Hans-Jürgen Schwab for their reliable and valuable input.

Last but not least, the GCEE would like to express his gratitude for the tireless efforts of its staff without which the German contribution to the report would not have been possible. Therefore, the GCEE specifically thanks Diplom-Volkswirtin and Diplom-Wirtschaftssinologin Ulrike Bechmann, Hasan Doluca, M.S., Dr. Malte Hübner, Dr. Anabell Kohlmeier, Dr. Heiko Peters, Dr. Stefan Ried, Diplom-Volkswirt Dominik Rumpf, Dr. Christoph Swonke, Dr. Marco Wagner and Dr. Benjamin Weigert. Special thanks go to Dr. Ulrich Klüh, whose input as Secretary-General until July 31 contributed considerably in preparing this report. Thanks also go to Dr. Jens Clausen, who as Secretary-General from August 1 on contributed to this report by coordinating the work of the staff and providing valuable inputs.

6. All views expressed in this report as well as all remaining errors should only be attributed to the authors mentioned below.

Paris and Wiesbaden on December 6, 2010

Conseil d'Analyse Économique

Christian de Boissieu Jean-Philippe Cotis

Michel Didier Christian Saint-Etienne

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung

Peter Bofinger Wolfgang Franz

Christoph M. Schmidt Beatrice Weder di Mauro Wolfgang Wiegard

Content

Charts

Tables

Boxes

CHAPTER I

Conceptual Foundations and Guiding Principles

Conceptual Foundations and Guiding Principles

1. As the world is emerging from its worst economic crisis of the last six decades, there is a broad consensus among policy makers and the general public that this should be a moment of pause and sincere reflection. From the vantage point of economics and statistics, three intimately related key questions should form the focus of such considerations. First, how can we improve our monitoring of **economic performance** in order to allow policy makers to gauge the current state of affairs and to react in a timely and appropriate fashion when crises emerge? Second, how can we broaden our perspective from its current focus on economic performance to an assessment of **quality of life** more generally in order to appreciate what really counts for human welfare? And third, how can we design warning signals alerting us whenever the current manner of organizing our lives endangers **sustainability** so as to correct our course of action for the sake of our own future and that of generations to come?

These are the questions which are addressed in this **joint study** by the French Council of Economic Advisors (CAE) and the German Council of Economic Experts (GCEE). The study has been conducted throughout the year 2010, at the request of the French President and the German Chancellor, and represents the result of inspiring and intense academic debates, detailed data work at both Councils and at the French (INSEE) and German (Destatis) National Statistical Offices, as well as a wide range of consultations of public officials, researchers and representatives of the many initiatives currently ongoing in the area of statistical reporting on human welfare. This study is **not exclusively** meant to be an **academic study**, venturing into the philosophical depths of assessing the state of mankind. Rather, despite its relentless intellectual aspiration, it **deliberately** sets out to represent a **pragmatic guide** to accounting for the current state of affairs. Taking as its point of departure the report of the "Commission on the Measurement of Economic Performance and Social Progress" (Stiglitz-Sen-Fitoussi Commission, SSFC), it discusses how comprehensiveness and accuracy might be traded off optimally against parsimony and cost so as to provide a reliable basis for regular, timely and digestible reporting on the three key issues regarding human welfare.

2. The first and arguably most important conclusion of our study is the **dismissal** of any single-indicator approach to measuring human progress as being insufficient. Life is simply too complex and the demands on statistical reporting are too diverse to allow a meaningful condensation of the current state of affairs into a **single comprehensive indicator**. While such a single indicator would emphasize parsimony and could be communicated easily, it could hardly do justice to the informational demands of modern democratic societies. Instead, we **suggest** that comprehensive statistical reporting should entail a **dashboard of indicators**. The dashboard we propose is meant to be open to discussion. It is rich enough to facilitate a meaningful discussion of the relevant facets of human welfare, yet it is not overwhelmingly extensive. Moreover, it provides a balanced representation of the three areas addressed by the key issues economic performance, quality of life and sustainability. This approach acknowledges that monitoring material well-being is an indispensable prerequisite for rational economic policy, that life is about more than material well-being but that human progress in non-material aspects is quite difficult to capture, and that it is wise to take a long-term perspective by outlining the consequences of unmodified human behaviour.

1. The challenge

3. At the end of the year 2010, when this report is being released, the world is slowly emerging from its worst economic crisis of the last six decades. This crisis has been an **extremely disruptive event**, shaking up the global economy, sparing hardly any region and eroding substantial parts of hard-won economic progress throughout the world. This chastening experience has challenged many preconceptions about the functioning of modern economies and about globalization. Specifically, social scientists, policy makers and the general public have been alerted to some **sobering insights**. First, given that mankind as a whole is richer than ever before in human history, it seems all the more intolerable that so many people are still being excluded from prosperity. And when economic disaster strikes, it would be desirable to be informed more swiftly and more precisely about emerging problems as they are unfolding and not – as it is the case now – after recession has already taken hold of the economy.

Second, while during economic boom periods it might be tempting to forget that market economies are characterized by fluctuations in growth, the crisis has brutally reminded us all that recessions and perhaps even depressions are a fact of life, historically as well as in the modern era. This aggregate **uncertainty** adds to the many possible sources of **instability** that may cast a shadow on individual lives, such as inequality, disease and persecution. Households and enterprises would surely be willing to sacrifice some average growth in gross domestic product (GDP) in exchange for lower volatility around the growth path, although it is difficult to determine how much.

More generally, in the wake of the crisis the realization is sinking in that life is about more than material well-being. Even an era which **superficially** might seem very **prosperous** can abruptly be revealed as falling short of its promises once the bubble bursts. Finally, material growth may occur at the cost of environmental degradation or otherwise endanger long-run economic stability. Whenever current economic success turns out to be **unsustainable**, it is not the original perpetrators but other people who will have to pick up the tab.

4. These insights could have severe **normative implications**. They have even prompted some people to believe that a lower rate of GDP growth is a necessary condition for humanity to comply with the objective of environmental sustainability. They argue that if **relentless pursuit** of economic growth is indeed endangering our resource base, this process cannot go on forever. And yet, when governments throughout the world were borrowing heavily to finance discretionary measures to counter the recent economic downturn, they were borrowing heavily from future generations. Consequently, if we were to deprive those generations of the opportunity to realize substantial economic growth, they would suffer. In that sense, economic growth is **crucially needed** to reduce unemployment, to improve well-being, to facilitate the "catch-up" process by developing economies, and to possibly ease disputes about the distribution of prosperity. But to win over its critics, it should concurrently seek to be smart growth which displays low carbon content and avoids negative side-effects on well-being. Clearly, the normative question of whether we should prioritize economic growth or other outcomes can never be discussed meaningfully on the basis of regular statistical reporting alone.

However, a discussion of such fundamental normative questions crucially presupposes the availability of timely, regular and accurate **statistical reporting** for deriving a meaningful portrait of the state of affairs. On its own the observation of current indicators of economic performance cannot clearly tell us whether an economy is on a desirable path or not. Therefore, the **positive questions** "Where do we stand?" and "Which state will our path lead us to?" are the first questions we should be able to answer regarding a wide range of relevant outcomes. Without a comprehensive statistical information system, we would not be able to address questions such as these, let alone conduct an informed normative debate. The desire to find a firm basis for this important debate explains the growing hunger for detailed and yet digestible information and, thus, for comprehensive statistical reporting that strikes the right balance between condensing the **essential facts** out of the wealth of data and retaining **enough detail** to do justice to the complexity of affairs.

Most importantly, the recipients of statistical information are **heterogeneous** themselves with respect to their preferences, capabilities and societal roles. Thus, although the objective of all statistical work is to reduce complexity, it needs to reflect the experiences of any modern democratic society, such as rapid structural change, technological progress and the forces of globalization. In a world as complex as ours, citizens and policy makers alike need detailed information so as to better understand what is important for their individual existence as well as for a collective good life. Well-informed citizens not only participate more actively in the democratic process, they also detect undesirable developments earlier. An important reason why the public has increasingly expressed dissatisfaction with the current focus on measures of economic performance is an **increasing gap** between the results of statistical reporting and individual perceptions of welfare. More specifically, although real income has risen in many countries throughout the last decades, self-reported well-being of the population has not increased in step (Easterlin, 1974; Frey and Stutzer, 2002).

5. Accordingly, measuring economic performance and well-being has again taken centre stage, both in the public and in the academic debate, building on a long tradition of this topic in economics and statistics. The relevance of this theme is evidenced, for example, by the European Commission's strategy for Europe 2020 with its demand for "smart", "sustainable" and "inclusive" growth. This is the point of departure for the **present study** whose goal is to provide the basis for timely, regular and accurate statistical reporting which strikes the right balance between comprehensiveness and parsimony. In this endeavour, we have been able to formulate our arguments by standing on the proverbial shoulders of giants. Specifically, in February 2008 the French President initiated the **SSFC Report** on measuring economic performance and social progress which was ultimately released in September 2009 and proved to be a **landmark study** in this debate. Hardly any other contribution to this area of research has stimulated such an intense discussion among policy makers and the general public. Moreover, since its release **numerous initiatives** have been taken or intensified, especially by statistical offices throughout Europe, to improve the state of statistical reporting further in areas as diverse as economic performance, non-material well-being and environmental sustainability. Building on years of work aimed at improving statistical reporting, many of these initiatives have been inspired directly or indirectly by the SSFC Report.

To many critical observers in the social sciences the SSFC Report **changed nothing**. After all, it addresses a subject that has been intensively discussed in economics and statistics for decades. Specifically, it has long been acknowledged by economists that GDP is not meant to be a gauge of human welfare, and that, taken by itself, it cannot satisfy the desire for information expressed by modern democratic societies. Nonetheless, GDP and its components have remained an indispensable guide for policy makers and the public. Without it, societies would be at a loss when it comes to judging short-term economic developments and the need for policy action. It is no coincidence that the foundations for GDP and the modern system of national accounts were laid in the 1920s and 1930s. Against the backdrop of severe economic dislocations not unlike the one we are facing today, the need for accurate measures of economic performance is obvious. The first formal national accounts were widely perceived as fulfilling this need. Since then, they have been steadily improved with a view to ensuring that they continue to be useful guides for economic policy. Thus, there are good reasons why GDP has been at the forefront of regular statistical reporting and economic policy – and these reasons are still relevant today.

To other observers the SSFC Report **changed everything**. While GDP as a concept remains indispensable, there are still areas in which improvements in measurement are required. As our societies and economies change, the focus of quantification also changes. For example, structural change required shifting the emphasis from the measurement of agricultural production to industrial production and services. Similarly, new modes of producing knowledge require new methods of measuring investment in this production. And yet, while the constant improvement of GDP measurement has always been a focus of statistical offices worldwide, the SSFC Report successfully alerted the general public to the importance and the intricacies of measuring societal progress and placed non-material aspects of well-being and the issue of sustainability at the top of the agenda.

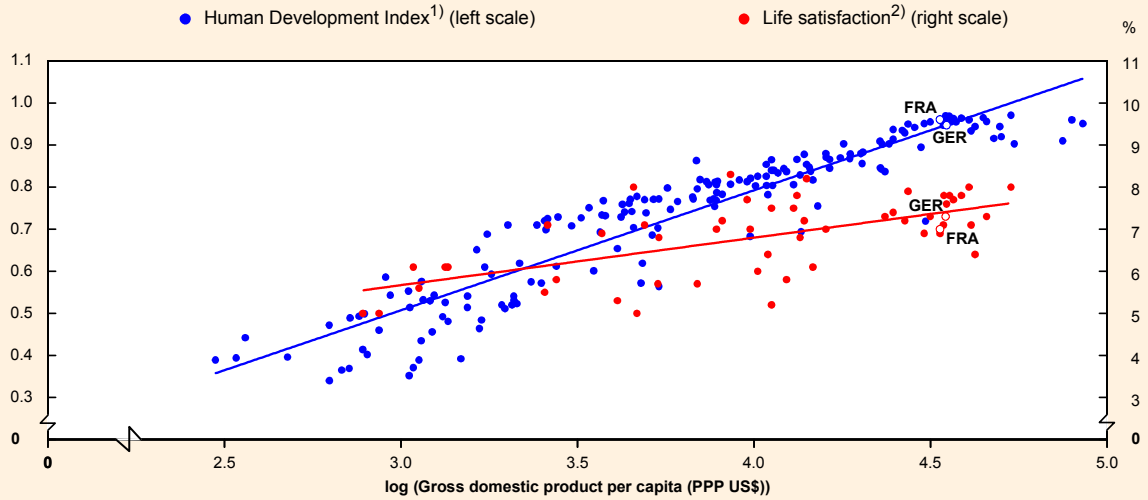
6. This positive perspective is our point of departure. The SSFC Report and the intense societal debate that followed in its wake provide a **tremendous opportunity** to move things forward. Most importantly, moving beyond GDP is not only a satisfying intellectual venture, it is also a **worthwhile endeavour**, notwithstanding the many obstacles that confront statistical work in the harsh reality of regular reporting. After all, we will never be able to "measure" societal progress perfectly. Rather, all we can hope for is to find indicators amenable to regular statistical reporting that allow an **approximation** to the true state of affairs. Inevitably, a margin of statistical uncertainty would always remain even if all systematic measurement errors could be excluded. It follows that all further work on improving statistical reporting by moving beyond GDP would be futile if the **correlation** between the standard measure of economic performance and alternative measures of societal progress turned out to be (close to) perfect.

Fortunately for our case, many survey-based measures of life satisfaction display significant, but **imperfect correlations** with GDP in cross-sections of countries and regions (Chart 1). While this pattern is less obvious over time within countries (Chart 2), there is some evidence confirming a high, albeit imperfect correlation within countries (Schmidt and Kassenböhmer,

2010). But moving forward is **not** at all **an easy task**. Most importantly, these high correlations imply that there is hope for enhancing statistical reporting meaningfully only if the **signal** that statistical research can **extract** from the raw data is highly informative. In operational terms, an indicator which is highly correlated with GDP is only able to add useful information if the statistical uncertainty associated with it is small. Unfortunately, more often than not, the intellectually most inspiring and original indicators of human welfare might not satisfy this requirement. And since our objective is to suggest improvements to regular and timely statistical reporting, this requirement is even more restrictive.

Chart 1

Correlation of gross domestic product and well-being



1) Human Development Report 2009.– 2) Mean of all answers; latest values between 2004 and 2008.

Data

Sources: UN, The World Bank, World Values Survey

Chart 2

Correlation of life satisfaction and gross domestic product

1) Mean of all answers.– 2) Gross domestic product per capita in constant prices.– 3) Until 1990: Western Germany.– a) Missing value.

Data

Sources: IMF, World Values Survey

7. The remainder of this chapter proceeds as follows. The second section briefly reviews the state of affairs after the release of the SSFC Report, while the third section outlines our own strategy for contributing to an improvement of regular statistical reporting. The fourth section summarizes our results, introducing a dashboard of indicators for the three key areas economic performance, quality of life and sustainability. As an illustration of the suggested dashboard we present results for France and Germany. Finally, we briefly sketch our perspective for the road ahead.

2. State of affairs

8. The history of statistical reporting on human welfare documents the fact that policy makers and the general public have been confronted for decades with **comprehensive sets** of economic and social indicators for short-term and medium-term policy analysis. Already in 1963, the law establishing the GCEE prescribed a multidimensional framework including price stability, a high level of employment, plus external balance with steady and appropriate growth, while additionally emphasizing the need to analyze the distribution of income and wealth. Similarly, since its creation in 1997, the CAE has traditionally been charged with monitoring and analyzing a wide range of issues, from distributional topics to questions of climate change. Notwithstanding these **traditions**, there has been an increasing awareness that the general public and policy makers alike have focused too much on GDP, and that other aspects of well-being have attracted insufficient attention.

It is therefore not surprising that recent years have seen a resurgence of work to improve our understanding of how societies are progressing, over time and relative to each other. In recent years numerous investigations and **initiatives** have been carried out at the Organisation for Economic Cooperation and Development (OECD), the United Nations, the World Bank, statistical offices and other organizations. As to academia, doubts with respect to whether GDP represents an accurate measure of all aspects of economic performance, let alone of well-being, can be traced back to Kuznets in the 1930s (Kuznets, 1934), one of the most prominent architects of national accounting. The debate was frequently revived over the last decades, for example by Nordhaus and Tobin in the 1970s (Nordhaus and Tobin, 1972). There is now a vast, well-surveyed literature on the measurement of well-being (Fleurbaey, 2009).

Most recently, the **SSFC Report** has led to an intense debate on the usefulness of a multitude of indicators for measuring economic performance and social progress, outlining three areas that any comprehensive statistical reporting should cover: economic performance, quality of life and sustainability. This intellectual framework emphasizes that the framing of timely and appropriate economic policy will always require the close monitoring of **economic performance**. It also acknowledges that a rich spectrum of facets of material and non-material well-being taken together affects the **quality of life** of individuals, families and households. And finally, it reminds us that severe disruptions might be looming ahead if we were to fail to take a **long-term perspective** regarding the consequences of unmodified human behaviour.

Elements of economic performance and social progress

9. A prerequisite for examining alternative approaches to measuring economic performance, quality of life and sustainability is an agreement on precisely what we intend to measure. The answer to this question is, of course, closely related to the **goals** that societies and, hence, policy makers want to pursue. Some light may be shed on this issue by asking the public and policy makers direct. The Euro-Barometer poll of 2008, for example, showed that more than two-thirds of EU citizens felt that social, environmental and economic indicators should be used equally to evaluate progress (European Commission, 2009). The **EU strategy 2020** currently underway also refers to three main objectives, namely "smart growth" (i.e., developing an economy based on knowledge and innovation), "sustainable growth" (i.e., promoting a more resource-efficient, greener and more competitive economy) and "inclusive growth" (i.e. a high-employment economy delivering economic, social, and territorial cohesion).

10. There is little doubt that economic policy makers need regular, timely and accurate indicators of **economic performance**. While the SSFC Report insightfully discusses a range of measurement problems concerning GDP, these arguably do not require special attention when it comes to designing short and medium-term economic policy. For the **short-term perspective** of macroeconomic policy with a time horizon of one to two years, GDP as an indicator of current value added still seems to be the most informative indicator of economic performance. And, of course, even within this area of economic policy, the research focus typically goes "beyond GDP" by analyzing data on unemployment, inflation, short-term business activity and consumer or business sentiment. For the purpose of taking a **medium-term perspective** the usefulness of GDP is somewhat more limited. But it again depends on the questions that policy makers want to address. While the shortcomings mentioned in the SSFC Report are relevant for an analysis of material well-being, they seem of much less relevance where issues concerning an economy's medium-term performance are concerned.

- GDP disregards the issue of depreciation, but the differences in both growth rates and levels between gross national income (GNI) and net national income are typically relatively small.
- It is also true that there are major differences in levels and growth rates between GDP and GNI due to payments and receipts of income to and from abroad. However, while this issue is relevant for an assessment of material well-being, it is quite irrelevant for an analysis of an economy's performance.
- International differences in GDP or GDP per capita reflect different preferences for goods and leisure. But for assessing economic performance one might also look at GDP per employed person or at GDP per hours worked.
- The SSFC Report correctly states that the level of non-market activities by households tends to differ among countries. Nevertheless, GDP-based measures provide an important picture of the performance of the market sector of an economy.

11. The clear distinction between economic performance and current material well-being is a highly commendable cornerstone of the SSFC Report. Measuring current **material well-being** is concerned mainly with the level, development and distribution of income, wealth, and consumption among households. Recently, public interest has shifted somewhat towards the issue of **distribution** and **inequality**, as is evidenced by the broad discussion about poverty rates and relative deprivation. Also, the increasing role of the public sector in providing certain services has to be taken into account. This refers not only to various types of services offered to individual households, for instance in the areas of health and education, but also to the supply of public goods more broadly. For Europe these aspects appear to be more important than for the US, for instance, as Europeans evidently tend to be less happy if inequality in society increases (Alesina et al., 2004). Generally, it is well known that GDP, being an indicator of production, has many flaws as an indicator of material well-being. They are mainly due to three factors.

- A given volume of GDP can be **distributed** in very different ways, across borders, between the private and the public sector, labour and capital and different income groups.
- A given volume of GDP can be **used** in very different ways, for consumption or investment purposes.
- And in addition to the market transactions recorded in the national accounts, there are important **non-market activities** by households that create material well-being.

12. One does not have to be very insightful to realize that not all that glitters is gold, and that all the material wealth in the world cannot buy a stairway to heaven. And yet, the SSFC Report raised this discussion to another level by devoting a whole chapter to a thorough discussion of this subject. Instead, **non-material** aspects of **well-being** ("quality of life") are equally important as the various elements of material well-being in generating human welfare. These non-material facets include, among others, health conditions, educational achievements, activities within the labour market, environmental aspects, social connections, political voice and security. Perhaps more than for other topics of social welfare, reporting on these facets requires an **implicit assessment** of individual and societal **preferences**. But as preferences tend to vary substantially among individuals and societies, it is questionable whether it makes sense to combine indicators of individual elements of quality of life into synthetic comprehensive indicators.

This caveat is even more disturbing when asking about ways to incorporate measures of subjective well-being ("happiness") into a standard corpus of indicators, an issue discussed in detail in chapter III. In particular, the implication of measuring happiness and well-being directly and inter-subjectively is, on the surface, that interpersonal comparisons become possible. However, one has to clearly point out the challenges to society if this were to be taken too seriously. What exactly happiness research can contribute to an understanding of progress is an issue that needs urgent attention.

13. The current economic crisis has demonstrated quite spectacularly that short-term gains might be eroded in the long-term, and should therefore be characterized as unsustainable if they rest on courses of action that create formidable imbalances. This holds for economic aspects, such as persistent and excessive borrowing by the private or the public sector or even whole countries, as well as outside the realm of traditional economic reasoning, most prominently in the area of ecology. In a very timely and insightful way, the SSFC Report invigorated the public debate about questions of sustainability which had been ongoing for almost four decades. Today, there is certainly a broad consensus among policy makers and the general public alike that it would be very fruitful to learn as early as possible about developments which might lead to disruptive corrections in the future, since life experience suggests that often there are two paths a society can go by, but in the long run, there's always time to change the road taken.

Guided by considerations regarding intergenerational fairness, the discussion of sustainability issues has a **long tradition** in economic thinking, relating to diverse aspects such as growth sustainability, environmental sustainability, fiscal sustainability and financial sustainability. For instance, the recent crisis has alerted economists and policy makers to the empirical regularity that high public debt in relation to GDP may negatively influence economic growth rates (Reinhart and Rogoff, 2010). But the appropriateness of **economic growth** per se has likewise been discussed for decades in academia and politics. To mention one early example, the law that established the GCEE in 1963 referred to "steady and appropriate growth" as a key objective, although environmental considerations were not considered to be such an important issue nearly fifty years ago.

Unresolved issues

14. While the **SSFC** has thoroughly analyzed the relevant issues and invigorated the debate on how to measure human welfare and social progress, it has left open a considerable number of unresolved issues.

- Which concrete set of **indicators** should, for individual countries or for groups of countries such as the European Union, form the core of the public discourse on progress, or, in other words, the new compass for the sovereign and its representatives?
- How can the objective of **refocusing** advocated by the SSFC be achieved in reality, given the understandable tendency of policy makers and the public alike to concentrate on highly aggregated, easily communicable measures?
- What precise **initiatives** have to be taken and publicly **funded** in order to resolve remaining conceptual issues and to gather missing data on a regular basis, taking into account the limited and perhaps even shrinking financial resources of statistical offices?

15. Since the release of the SSFC Report, many institutions have taken up the challenge of answering the questions that have been raised by it. Its proposals have led to intensive research efforts, above all in statistical offices, but also in other international and national institutions. As many proposals are related to questions of how to improve the statistical meas-

urement of economic performance and well-being, it is not surprising that mainly **statisticians** are engaged in these efforts to implement or refine the results of the SSFC Report. Moreover, while the follow-up work on the SSFC Report is a global project involving a diverse set of contributors, **European institutions** have played a prominent role. In the context of the "GDP and beyond" initiative, a concrete work plan has been pursued (European Commission, 2009), and this work will certainly be integrated into the Europe 2020 strategy.

As there are so many competent international and national statistical institutions involved in this debate, the focus of our contribution is guided by what we view as the **comparative advantages** of the joint research team formed by the CAE and the GCEE. Clearly, the natural realm of economics is the market process, which attaches prices to goods and services. Together with the assumption that market prices are shaped to a considerable extent by the wishes and desires which consumers associate with a certain item, this allows economists and statisticians to aggregate units of widely differing goods and all kinds of services to GDP, and use this construct as an indicator of economic performance and perhaps – under some qualifying assumptions – even well-being. Nevertheless, the convincing postulate that one needs to go "beyond GDP" leads to areas outside of this standard of economic inference. This applies above all to the measurement of non-material well-being.

16. Even so, there are several counts on which economists can contribute fruitfully to the quest for improvements in statistical reporting on economic performance and well-being. First, the search for a better measurement of economic performance is necessarily predicated on the need for **economic expertise**. Second, although economics is mainly focused on material items, its intellectual discourse rests on a set of tools that facilitate a wider panorama than this relatively narrow view. The **concept of utility** can be applied in ways that reach far beyond the utility that can be derived from the consumption of goods. Modern welfare economics and the approaches that build on it have broadened our horizon substantially. Happiness research is a relatively new area to which economists have contributed heavily. Third, and perhaps most important, there is virtue in **realizing** one's own **limitations**. Since economists are traditionally very cautious about interpersonal comparisons of welfare, they bring a sceptical approach to the debate that is desperately needed when it comes to interpreting statistical indicators. Spelling out the precise conditions which have to be satisfied to allow the aggregation of individual indicators into a summary indicator or an international comparison of numerical values, for instance, is an important prerequisite for any meaningful discussion of these issues.

Fourth, in economics the issue of sustainability tends to be addressed in an encompassing fashion that includes, but is not limited to, **ecological** sustainability. Indeed, economists have paid a lot of attention to the sustainability of **public finances**, and especially after the recent financial crises, the sustainability of private-sector **financial balances** has also moved to the fore of the economic research agenda. Finally, as economic thinking revolves around the concept of scarcity, economists are trained to identify and discuss **trade-offs**. In the context of regular statistical reporting on human welfare, the requirement of striking an ideal balance between comprehensiveness and parsimony is absolutely central.

3. Principles and obstacles

17. This study builds on a wide spectrum of existing and well-matured statistical reporting and recent initiatives which have generated a **large reservoir** of arguments, procedures and indicators regarding the monitoring of economic performance and well-being. The objective of our work is to discuss how comprehensiveness and accuracy in statistical reporting on human welfare should ideally be **traded off** against parsimony and cost in order to provide the basis for regular, timely and digestible information on the state of affairs. Inspired by the SSFC Report and the intense debate that followed in its wake, we organize our ideas according to **three areas** of application: material well-being, quality of life and sustainability. While we intend to produce a pragmatic guide to accounting for human welfare and, thus, always keep an eye on cost and on the advantages of utilizing statistical work that has already been implemented successfully, we feel free to suggest the collection of additional information whenever necessary to ascertain a satisfactory degree of comprehensiveness.

18. Our **main contribution** to the debate is threefold. First, we propose a **concrete set of indicators** that are relevant for short-term and medium-term policy decisions, taking into account the trade-off between the required comprehensiveness of an indicator set and its relevance for decision makers. In the end, concrete policies can only be based on a detailed study of all relevant information and on a broad list of indicators that mirror all possible facets of economic and social life. For the overarching policy-making process, however, it is essential to focus on a relatively limited number of indicators even if this implies that some aspects have to be omitted.

Second, we propose a concrete way of communicating these indicators to the public by devising a **dashboard** based on the **three pillars** that logically follow from the main themes of the SSFC Report.

- The first pillar includes indicators for assessing **economic performance** and current **material well-being.** It is based primarily on the economic flows that are recorded in the systems of national accounts and the statistical data on the distribution of income.
- The second pillar focuses on the non-material aspects of well-being and proposes separate indicators for a set of well-defined dimensions of the **quality of life**.
- The third pillar is devoted to the issues of **sustainability**, i.e. the question "whether we can hope to see the current level of well-being at least maintained for future periods or future generations" (Stiglitz et al., 2009).

This three-pillar approach allows a comprehensive assessment of a country's economic performance and well-being over time as well as comparisons with other countries. It is important that the information provided by each pillar is not used in isolation, but that the pillars are used simultaneously with regard to all three dimensions. This would facilitate a discourse on policy that accounts for the trade-offs between different areas of well-being as well as between short, medium and long-term concerns.

Third, in our treatment of sustainability, we not only focus on environmental concerns but also include issues of **economic sustainability**. Needless to say, this encompassing view of sustainability is in no way intended to belittle the importance of the other main dimensions of sustainability, namely the preservation of ecological capital and the existence of sufficient social and political capital. Conversely, public investment in the ecological modernization of our economies and in social cohesion is only possible on the basis of stable government finances and a private financial sector that is spared from disruptive corrections.

19. The intellectual construct of our proposed dashboard and the systematic organization of ideas and arguments applying to each of the three areas of application are only the starting point. Most of the effort has to be spent on the detailed choice of dashboard entries from the – sometimes abundantly rich and in other instances frustratingly meager – reservoir of candidate indicators. This choice requires a detailed assessment of **statistical indicators** as regards their **quality**. The quality of statistical indicators is sometimes described as depending on the following three crucial criteria: relevance, consistency with theory, and measurability.

- The requirement for "**relevance**" is quite obvious. It implies that indicators should be chosen such that they adequately capture changes in the current or future level of a given aspect of well-being.
- "**Consistency**" implies that indicators are designed in accordance with theoretical considerations. Specifically, these considerations require congruence of a measurement with a relevant dimension of well-being. In addition, they define the limit for aggregating heterogeneous information. Finally, they force the researcher to take into account that some aspects of well-being will remain unobserved, so that the best one can hope for is to identify latent variables or proxies.
- "**Measurability**" implies that indicators are actually observable and raw data can be collected at reasonable fiscal and private cost and without violating data protection and privacy laws. Each data collection incurs direct and indirect costs. Direct costs for a survey result, for instance, from the salary of the interviewers and expenditure on data processing machines. Indirect costs result from the opportunity cost of interviewed companies or citizens who are not compensated for the time they spend answering the questionnaire.

20. While the main objective of our report is to identify a limited set of indicators that are both politically relevant and capture the main dimensions of current and future well-being, methodological consistency and measurability are constraints that have to be taken seriously. In this respect, the crucial aspect is **cost considerations**. Given unlimited resources, most methodological barriers and data limitations could be resolved, at least in principle. However, scarce resources are to be taken into account, as statistical offices face tighter budgets and our societies prepare for the pressing task of consolidating public finances.

It is very difficult to obtain an exact cost estimate for each of the three pillars of our proposed dashboard. At one end of the scale, there are no extra costs of using already available data. In our considerations, we are intentionally **pragmatic** and give priority to using existing indica-

tors over requesting new data collection and indicator construction whenever possible. At the other end of the scale, the exacting demands which the indicators in the suggested dashboard have to satisfy leave no other option than either to collect additional information or to collect existing information more frequently. In between, there is a wide spectrum of requirements for additional work and corresponding additional cost. In cases where the quality of already collected indicators is improved further and concepts are harmonized internationally, it is easy to underestimate the cost involved. As all the national statistical offices of the European Union and Eurostat are currently working to improve the official statistics further, a reliable cost estimate might be extracted from their cost reviews.

21. Notwithstanding the aforementioned constraints, using the second pillar of our dashboard, quality of life, as an example, we are able to provide a very **rough indication** of the possible cost involved. In this application, the collection of new information by new surveys, the addition of further questions to existing surveys and improvements in the quality of surveys by extending the number of interviewees are natural desires emerging from the discussion of various non-material facets of well-being. These requests might be difficult to refuse, even though we try to be as pragmatic as possible and always gauge very carefully how we could utilize the reservoir of existing work for selecting a useful dashboard entry. Even so, as the detailed discussion on quality-of-life issues demonstrates very clearly, when we confront existing indicators with the exacting demands that we adamantly must raise in order to construct a meaningful dashboard, we are not always successful in finding a suitable ready-made candidate indicator. In that case, there are just two options: either to omit the desirable entry from the dashboard – since no entry is arguably much better than a misleading entry – or to bear the additional cost and fund the additional data collection.

The cost of any additional data collection varies directly with the scope of the survey, the number of respondents being interviewed and the frequency of interviews. This generates three-dimensional optimization problems which have to be solved one at a time, and not once and for all. For example, preparing and implementing a completely **new survey** like the Programme for International Assessment of Adult Competencies (PIACC) costs about 10 million euro for the first wave in Germany. If the PIACC were to be repeated on a yearly basis, the cost per wave would be substantially reduced. The cost of conducting the EU-Statistics on Income and Living Conditions (SILC) for Germany is around 3 million euro for about 75 survey questions for the last available cost calculations of the year 2006. Thus, a simple approximation of the cost of adding a **further question** to the EU-SILC questionnaire could be inferred from the average cost of one question, leading to an estimate of about 40,000 euro for adding a question to EU-SILC. Finally, improving the quality of indicators by increasing the number of **interviewees** in a survey is costly as well. For example, the cost of increasing the number of people interviewed by 28,000 to 30,000 would cost around 3 million euro. This might be worth it, though, given our ambitious objective.

4. Key results

22. This report is aimed at providing input for the much broader debate on how to measure the state of human welfare. It deliberately delineates its arguments according to **three areas**

of application, namely material well-being, quality of life and sustainability. Each of these areas is discussed in detail in one of the following substantive chapters. This section provides a brief executive summary of the arguments and the concrete suggestions that emerge from the detailed discussions. In addition, to illustrate our suggestions we present an **application** of our proposed dashboard to the case of France and Germany. We separate the dashboard into the three sections associated with the three areas of application.

As many of the entries of our proposed dashboard are not amenable to international comparison by construction, this **illustration** should not be mistaken as a serious gauge of the relative economic performance of these two countries, let alone of the quality of life of their citizens vis-à-vis one another. As is explained carefully in the respective substantive chapters, most entries have to be viewed as providing information on the developments within an economy over time but are rather uninformative on cross-country comparisons at a single point in time. The dashboard included here should therefore simply be read as evidence that our work seeks to help to take the discussion on the statistical reporting on human welfare and societal progress one step further.

Economic performance and material well-being

23. In the first area of application, we distinguish the monitoring of economic performance from an assessment of material well-being. We start from the insight that GDP aggregates the **value added** of all market activities and of input-related measures of government activities – nothing more but also nothing less. While GDP is well-suited for this purpose, it is by no means flawless. Specifically, we discuss measurement problems such as the difficulties generated by capturing the intensity of economic activity in the **service sector**, in particular, when these services are provided by the public sector. Another problem that is addressed in more detail is the existing focus on market production, which leads to the omission of **non-market** economic activities such as the household production of services. The extent to which we can base this discussion on past work is evidenced by the fact that the German Statistical Office already reported information on household production in the time use surveys of the year 1991/92 and 2001/02 (Statistisches Bundesamt, 2003; Schäfer, 2004). We also contemplate whether and how to include economic activities occurring in the **shadow economy** into regular statistical reporting.

Furthermore, our discussion acknowledges that GDP, while being correlated with many variables that are crucial for well-being, is not per se a perfect metric of **well-being** (Costanza et al., 2009). More generally, all measures of progress relying on market prices are only reliable gauges of well-being in the absence of serious **externalities**. In addition, as GDP might vary according to a society's **preferences** for work and leisure, it has to be asked how these varying preferences can possibly be accounted for in statistical reporting. Finally, as an aggregate measure, GDP tends to conceal **distributional outcomes**. As emphasized by the European Commission, “social and economic cohesion are the overarching objectives of the Community. The aim is to reduce disparities between regions and social groups.” (European Commission, 2009). Growth accounting based on GDP ignores income disparities. Wealth and its distribution are not captured by GDP at all.

24. Correspondingly, the second chapter reviews the five first recommendations of the SSFC Report. Its first recommendation is the request to assess the current state of **material well-being** on the basis of income per capita and consumption rather than GDP which, as outlined above, nevertheless remains a valuable indicator of economic performance. Second, the SSFC Report recommends emphasizing the household perspective when material well-being is at issue, while the third recommendation alerts researchers to consider wealth as an important facet of material well-being. A fourth recommendation of the SSFC Report emphasizes the importance of distributional characteristics of income, consumption and wealth, and, finally, a fifth recommendation suggests broadening the perspective to include non-market activities.

Our discussion has been informed by the insight that, although there is always scope for augmenting material well-being further, for **wealthy societies** such as those of France and Germany, it is already an achievement to maintain the existing high level of productive activities. Thus, monitoring economic performance remains an important task and implementing refinements of GDP which serve this task even better is an important objective for economic and statistical research. Nevertheless, the SSFC Report reminds us of the need to be aware of the limitations of GDP as a measure of well-being, a theme which – as we have emphasized throughout this chapter – has been discussed by economists for many decades. Therefore, our report explores promising avenues for proceeding from the measurement of economic performance to an assessment of material well-being.

25. Most decision-makers would certainly like economists to provide them with "the" ultimate indicator of material well-being. We fully agree with the overarching conclusion emerging from the SSFC Report that this idea is totally unrealistic. In order to proceed from this fundamental insight towards the practical implementation of more realistic alternatives to traditional statistical reporting, we propose **six indicators** that seek to strike an appropriate balance between comprehensiveness regarding economic performance and the current state of material well-being, on the one hand, and parsimony, on the other. These indicators are:

- GDP per capita,
- GDP per hours worked as a measure of economic productivity,
- employment rate for the 15-64 age group,
- net national income per capita,
- final consumption expenditure per capita, including government consumption,
- an internationally harmonized distribution measure of net income per consumption unit (income quintile share ratio S80/S20).

26. We have also proposed **concrete steps** that need to be taken rapidly – notably the harmonization of panel data on household income – to facilitate consistent measurement of changes in income distribution, such as the EU-SILC (Survey on Income and Living Conditions) panel. In particular, the sample size should be expanded if we want to gain more com-

prehensive knowledge not only of differences in income distribution but also of other factors linked to well-being. Regular studies comparing time use across countries should also be considered. Furthermore, we have outlined the need for further statistical advances in fields such as in-kind services and intangible activities – and, more generally, in the statistical coverage of various economic sectors.

Reforming the system of indicators of economic performance and current material well-being is important. But to effectively develop a new compass for policy-making, the crucial step will be to anchor communication on progress to a system of indicators that takes better account of non-material sources of well-being and the sustainability of current modes of behaviour and levels of well-being. These issues are addressed in the two other substantive chapters of this report.

27. The part of the dashboard that constitutes an **indicator set** for **material well-being** yields the following observations for France and Germany (Table 1). GDP per capita in both France and Germany is substantially above the EU 27 average and has increased in both countries over the period from 1999 to 2009, although not as strongly as the EU 27 average. Similar observations hold for the indicators GDP per hours worked and final consumption expenditure per capita. The increase in France was stronger, however, than that of the EU 27 average. The employment rate increased particularly strongly in Germany in the last decade. Net national income per capita is correlated with GDP per capita and is useful when comparing France and Germany to countries with large cross-border factor income or investment flows. The income quintile share ratios S80/S20 indicate that in Germany the ratio between income received by the top quintile and the bottom quintile is 4.8 in 2008 (and has increased relative to 2000) and 4.2 in France in 2008 (as in 2000).

Table 1

An Indicator Set for Material Well-Being

Indicator	Unit	Last observation (2009)			Change over ten years (1999 - 2009)[1]		
		Germany	France	EU 27	Germany	France	EU 27
GDP per capita	€	29,278	29,571	23,588	+ 1.8	+ 2.7	+ 2.8
GDP per hours worked[2]	€	43.2	48.3	32.8	+ 2.4	+ 3.3	+ 3.2
Employment rate[3]	%	70.9	64.2	64.6	+ 5.7	+ 3.3	+ 2.1
Net national income per capita	€	25,220	25,586	.	+ 2.0	+ 2.4	.
Final consumption expenditure per capita	€	23,001	24,538	19,017	+ 1.9	+ 3.3	+ 3.1
Income quintile share ratio (S80/S20)[1)4)]	.	4.8	4.2	5.0	+ 1.3	+ 0.0	+ 0.1

1) Annual average growth rate except employment rate and income quintile share ratio.– 2) Between 2000 and 2008.– 3) Population aged between 15 to 64 years.– 4) Ratio between income of the top quintile and the lowest quintile. EU 27 between 2005 and 2008.

Data

Sources: EU, OECD

Quality of life

28. There are a myriad **social indicators** most of which have their merits per se. In France and Germany, statistical offices regularly publish a considerable number of figures concerning health, education, security, and other non-material aspects of well-being. Moreover, various research programmes in social sciences offer a wide range of indicators about subjective well-being. In Germany, research on measuring social welfare by – both objective and subjective – social indicators has a long tradition, as is evidenced by publications by GESIS-ZUMA and Wissenschaftszentrum Berlin für Sozialforschung (WZB) (Statistisches Bundesamt et al., 2008; GESIS-ZUMA, 2007). In France, since the 1970s a report "Données sociales: La société française" has been published regularly. Actually, it comes as a slight surprise that the SSFC Report omits mentioning these achievements.

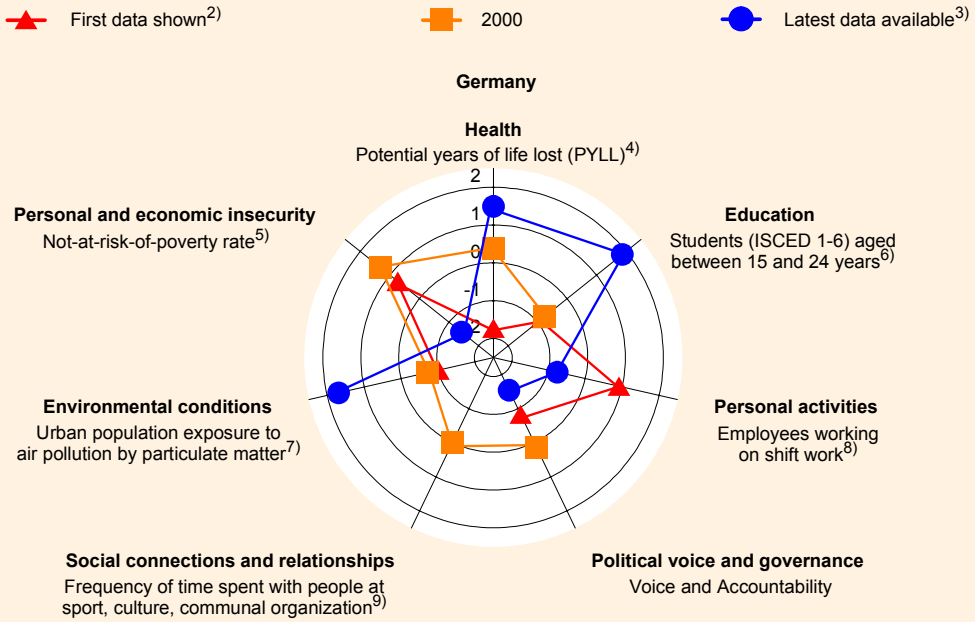
Therefore, possible reservations against the usefulness of social indicators cannot stem from a lack of information. Rather, the challenge is more how to use this **plethora of information** appropriately and how to improve the international comparability of quality-of-life indicators. These methodological issues are discussed thoroughly by Fleurbaey (2009). Two major issues arise. First, preferences among people differ even within one country. Hence, it is not clear what comparisons between **subjective assessments** of people about their well-being, let alone happiness, really mean. Such comparisons are even more doubtful if people care about each other or about their relative standing in society.

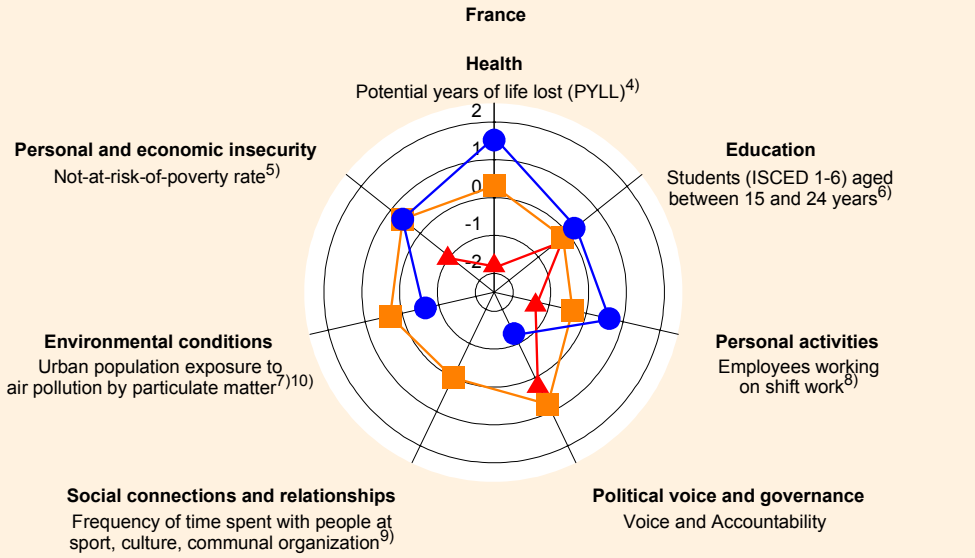
This caveat particularly challenges happiness research, despite the considerable methodological progress in this research area that has taken place in the past years (Frey, 2008; Layard, 2005). Consequently, these approaches unfold their potential mainly by providing information on whether the same people are better off or not, i.e. the sign of a **change** in well-being. These caveats carry over to synthetic social indicators like the Human Development Index (HDI). In addition, the weights of the various domain indicators are up for discussion. How might various indicators be traded off against each other, such as, for instance, the suicide rate against the literacy rate (Fleurbaey, 2009). Therefore it seems very sensible to opt for presenting life's complexities for what they are, rather than giving priority to utmost parsimony.

29. Acknowledging this intricate discussion, the third chapter of this study lays the ground for enhanced regular reporting of the state of well-being that comprehensively covers a wide spectrum of facets of human existence. Regarding the results for quality of life, besides a summary of the most recent developments, the complexity of the matter requires that the bald numbers must always be elucidated and interpreted **carefully.** After all, the very nature of the various non-material dimensions of quality of life means that even the best indicators of the state of affairs are only **imperfect proxies** and should be discussed with all due consideration of their potential and their limitations before formulating any recommendation for policy action. Furthermore, we propose to visualize the results in the form of a radar chart which illustrates the developments along the seven dimensions over time and demonstrates the multifaceted nature of the phenomenon under study (Chart 3). But one should **never** fall into the seductive trap of constructing an **encompassing quality-of-life indicator** or surface measure, as easy as that might be in terms of calculation.

Chart 3

Non-material quality-of-life indicators[1)]





1) Own calculations; values are not comparable across countries. Average = 0; value higher than 0 implies better conditions and vice versa.– 2) Health: 1991, Personal activities: 1992, Political voice and governance: 1996, Education: Germany: 1992, France: 1993, Enviromental conditions: Germany: 1999, France: 2001, Personal and economic insecurity: Germany: 1992, France: 1995 .– 3) Health: 2006, Education and Personal activities: 2009, Political voice and governance and Enviromental conditions: 2008; Personal and economic insecurity: Germany: 2009, France: 2008.– 4) PYLL is a summary measure of premature mortality which provides an explicit way of weighting deaths occuring at younger ages, which are, a priori, preventable. In relation to 100,000 population, calculated by the OECD Secretariat based on age-specific death statistics provided by the World Health Organization.– 5) One minus share of persons with an equivalised disposable income below the risk-of-poverty threshold, which is set at 60 % of the national median equivalised disposable income after social transfers.– 6) In relation to the population in the same age group.– 7) The indicator shows the population weighted annual mean concentration of particulate matter at urban background stations in agglomerations.– 8) As a percentage of total employees.– 9) Only data available: 1999.– 10) For 2000: 2001 data.

Sources for calculations: EU, OECD, SOEP, The World Bank, World Values Survey

Data

30. The SSFC Report makes five recommendations with respect to quality of life, leaving it to further research to set the adequate **priorities** between them. First, measurement along all dimensions should be improved, with particular efforts necessary for social connections and relationships, political voice and governance, and insecurity. Second and third, inequalities should be assessed and interrelations between the dimensions explored. Fourth, various forms of aggregation should be made possible through adequate provision of information. And finally, subjective measures of well-being should be surveyed by statistical offices. As they are quite general, the CAE and the GCEE naturally agree with all of these five uncontroversial recommendations. In our own contribution, we have decided to improve the state of play regarding two areas touched upon by the recommendations with the objective of forming a solid basis for the actual application of the conceptual ideas.

The first contribution we make is with respect to **aggregation**. The construction of composite indicators is more than a mere technical issue, since it always involves a large range of serious identification assumptions. Our detailed discussion of this matter has led to the formulation of a pragmatic and yet, we believe, conceptually sound strategy. While we are adamant that aggregation across the dimensions of quality of life would have to rely on overly strong identification assumptions, aggregation within one dimension might be less controversial. Of the various methods available to aggregate within dimensions, we assess the potential of two of these methods to condense information. Furthermore, our discussion pays considerable attention to the **communication** of the results. In particular, we propose the publication of graphs that visualize the results.

The second contribution relates to concrete steps towards improving **measurement**. At first glance, measures of the dimensions of quality of life are in abundant supply. Some of its elements – mortality tables, violent crime – even belong to the oldest statistics collected regularly. Yet closer inspection reveals the imperfect state of affairs, as our detailed discussion has documented. Given the intensity of efforts spent by governments and statistical offices on this matter, however, there is ample reason to hope for rapid improvement. To improve the current state of affairs, one has to survey the existing measures within **each dimension** and single out the most important deficiencies. Major topics in this context are international availability and comparability, both between France and Germany and within Europe, and the frequency with which the measures are currently calculated.

31. Our discussion of these issues suggests that one does not have to leave the vantage point of economics to realize that life has more to offer than its material aspects. **Non-material** elements of **well-being** play an important role in determining individual fulfillment and satisfaction and societal progress. The third chapter discusses the difficult task of gauging non-material well-being at the individual level and, via the aggregation of individual information, at the level of societies. Moreover, it provides a first application of the **empirical strategy** emerging from this discussion to the cases of France and Germany, guided by the clear understanding that this analysis is a first step and not an end in itself. In this endeavour, we have made a series of **deliberate choices**, both at the conceptual and the applied level, balancing the desirable with the achievable.

Regarding the **conceptual discussion**, we strongly advocate what we have termed a **bottom-up approach**. We could start our search for a better grasp of the state of non-material well-being from survey information on individual "happiness", but fundamental questions of measurability and the risk that such inherently imperfectly defined measures of human satisfaction could too easily be manipulated into showing politically desirable results prevent us from embracing this approach. Instead, our advice would be to condense the ample information on diverse elements of non-material well-being as much as possible so as to make the information digestible by its recipients while simultaneously retaining as much of its complexity as necessary to reflect its variegated nature.

Our concrete empirical strategy starts from the definition of a range of **dimensions** that should not be aggregated any further in order to adequately capture life's complexity. In our application, we have been guided by the SSFC Report into choosing seven dimensions, some of which pertain to individuals themselves, such as health and education, while others describe the societal and physical context experienced by individuals, such as social connections and relationships and environmental conditions. The strategy then proceeds from dimension to dimension, one at a time, and identifies for each of them a series of **individual indicators** that capture its facets as comprehensively as possible. Finally, for each dimension separately, we select one **headline indicator** out of this reservoir to represent the dimension as optimally as possible. Whenever feasible, we engage in a procedure of statistical complexity reduction in order to cross-check our selection of headline indicators. Most importantly, throughout our analysis we have worked under the **constraint** that the indicators chosen need to be regularly available in order to facilitate a perpetuation of this report in future years.

32. The application of this strategy to two countries, France and Germany, has uncovered a set of results that are plausible in that they paint a **mixed portrait** of **societal progress** over the last decade. In particular, progress in terms of health, education (with some reservations), and environmental conditions appears to be highly congruent with the steady growth experienced in material well-being. And yet, while they are admittedly difficult to capture, the recent developments in other dimensions of non-material well-being, such as personal activities and personal insecurity, indicate that societal progress has not been achieved unequivocally across all relevant dimensions.

Sustainability

33. The fourth chapter starts from the insight that, although current economic performance and well-being might appear quite satisfactory, current paths of action, if persistently continued into the future, might well turn out to be unsustainable. In that case, they might require sharp and painful adjustments and perhaps even cause socially costly crises. One section of this chapter is particularly concerned with two facets of **economic sustainability**, growth sustainability on the one hand and external and fiscal sustainability on the other. Another section of this chapter addresses a third facet of economic sustainability, namely private sector financial sustainability. Throughout these sections, the discussion focuses on the **medium-term** and the **long-term** perspective.

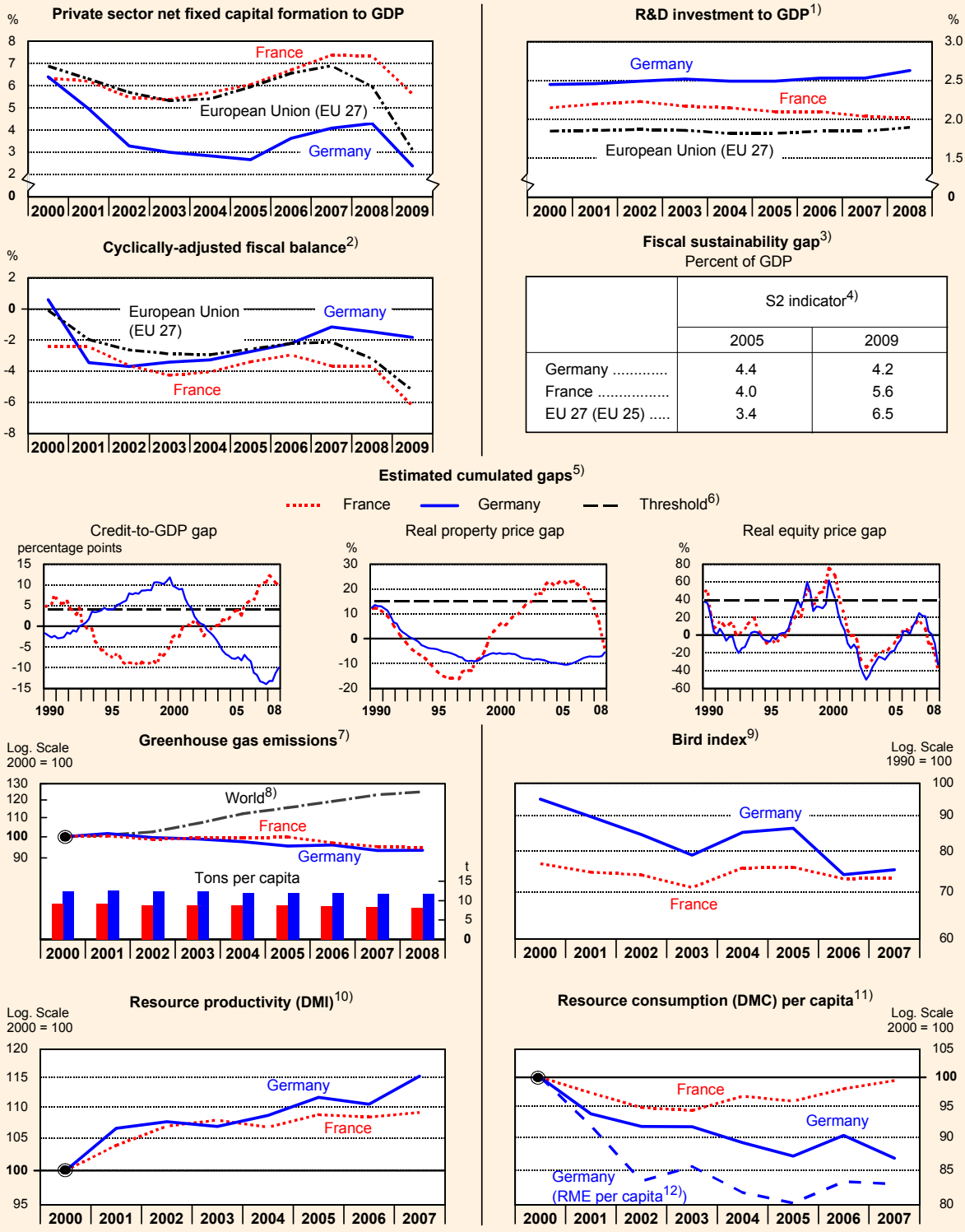
The first aspect of economic sustainability that we address in our analysis is the issue of **growth sustainability**. Specifically, we consider growth to be sustainable if a sufficient part of wealth creation in the economy is allocated to investment, irrespective of whether it is invested in material or immaterial capacities. Consequently, in order to emphasize the importance of capital accumulation for economic growth, we incorporate the ratio of **net fixed capital formation of the private sector** to GDP into our dashboard (Chart 4). While France's ratio has been moving closely with the EU 27 average, Germany's investment ratio has been below that of France and the EU 27 average from 2001 onwards. Moreover, since we require a reliable predictor of future overall productivity, and of expected trends in science, technology and innovation, we have chosen as a second indicator of growth sustainability for our dashboard **R&D investment** of an economy relative to GDP. Here, both Germany and France display a ratio that lies consistently above the EU 27 average.

The second aspect of economic sustainability, external and **fiscal sustainability**, is intimately related to the intertemporal budget constraint which is necessarily binding in the long term. Due to its inherent long-term perspective, this issue is also closely linked to concerns of intergenerational equity. Ultimately, when unsustainable fiscal and external positions have to be unwound, this can have painful consequences. As our concrete indicators of fiscal sustainability, we have chosen, first, the **cyclically adjusted public sector balance**, which, according to the Golden Rule of Public Finance, should not exceed net public investment. However, the cyclically adjusted deficit exceeded public net investment in both Germany and France in the period between 2001 and 2009. And as a second indicator of fiscal sustainability we have selected the **fiscal sustainability gap** as represented by the indicator „S2" in the European Commission's Sustainability Reports. To signal fiscal sustainability, this indicator should be negative or zero. For France, the S2 indicator displayed an adjustment need of 5.6 percentage points in 2009. For Germany, the S2 indicator documented an adjustment need of 4.2 percentage points (Chart 4). Given a positive sustainability gap, the indicator should at least decrease over time and eventually converge towards zero to ensure that current fiscal policies are sustainable.

34. The fourth chapter also discusses possibilities to augment the monitoring of current economic performance and well-being, which is regularly conducted by the statistical offices, by a complementary documentation of the state of **financial sustainability**. To this end, it suggests a set of indicators that signal unsustainable developments in the private and in the financial sector. Their objective is exclusively to investigate **excessive** fundamental and undesirable developments that are likely to lead to severe economic crises. While this objective is ambitious, the discussion makes it clear that it will never be possible to predict financial crises with certainty. What is offered here, though, is a small set of reasonably robust **early-warning** indicators that could alert policy makers and the general public in the event of fundamental undesirable developments in the financial sector. They are intended to be simple and manageable indicators for policy makers and the wider public who do not have the time and expertise to consider a plethora of disaggregated indicators or to employ stress testing or comprehensive early-warning models themselves.

Chart 4

Sustainability indicators



Fiscal sustainability gap3)
Percent of GDP

	S2 indicator4)	
	2005	2009
Germany	4.4	4.2
France	4.0	5.6
EU 27 (EU 25)	3.4	6.5

1) Gross domestic expenditure on R&D as a percentage of GDP.– 2) Source: EU. Percent of potential GDP.– 3) Source: European Commission „Sustainability Report 2009" and „ Long-term sustainability of public finances in the European Union" (2006).– 4) Ne-cessary adjustment of structural primary balance required to close sustainability gap.– 5) For details on methodology see Borio and Drehmann (2009a).– 6) The threshold is 4 percentage points for credit-to-GDP gap; 15% for real property price gap and 40% for real equity price gap.– 7) The annual greenhouse gas (GHG) emissions are estimated and reported under the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol and the Decision 280/2004/EC.– 8) Source: IEA and OECD. Only CO_2 emissions from fuel combustion.– 9) Source: EU. Only farmland birds.– 10) Ratio of Gross Domestic Product to DMI; DMI: Direct Material Input (abiotic materials which are directly used in the economy; materials used domestic extraction and physical imports).– 11) DMC: Domestic Material Consumption (total of all abiotic materials used up domes-tically; DMC = DMI – exports).– 12) DMC in raw material equivalents (RME).

Data

Despite these reservations, the three proposed indicators are – in our opinion – the most reasonable extract of the empirical literature concerned with the issue of leading indicators. We suggest looking at the ratio of total **private credit** relative to GDP, and at **real equity prices** as well as **real property prices**, both deflated by the consumer price index. More specifically, when monitoring these variables, we propose studying the cumulative deviation from their trends (their cumulative gaps). This suggestion can be implemented directly. Data on total private credit and equity prices are provided by national central banks, and data on property prices are collected by the Bank of International Settlements (BIS) and can be retrieved from the BIS. Of these three indicators, only one is currently showing warning signs: the credit gap in France (Chart 4). The data, however, end in 2008 and the credit-to-GDP gap is gradually receding. While this limited set of indicators should obviously not be understood as a substitute for detailed macro-prudential supervision or existing early-warning systems used by experts and sovereign authorities, their promise is to identify those economic developments early in the process which, if left uncorrected, might lead to stress situations. If these indicators signal an alarming development, policy makers should consult experts and authorities and if necessary take remedial action.

Regarding further work on this issue, especially on the supranational level, it is necessary to ensure data quality. Most vitally, there is a need for **harmonization** and **standardization** of data collection processes across countries to generate reliable and comparable information. This is all the more important as globalization in general and financial integration in particular force us to act on the EU-level – thereby involving 27 nation states. As harmonization is primarily a matter of setting standards for definitions, data collection processes and data quality, this should be a very cost-efficient but simultaneously valuable contribution.

35. Finally, but certainly not least importantly, the fourth chapter contains a detailed discussion on the statistical reporting on **environmental sustainability**. According to the current state of knowledge, rising levels of carbon dioxide and other greenhouse gases (GHG) in the atmosphere have already caused global warming, and will induce climate change on an even broader scale. Climate change has the potential to trigger major social and economic crises. Accordingly, GHG emissions should be a component of our dashboard. Of course, the figure that is most relevant for climate change is the **level** of **GHG emissions**. But climate change is a global phenomenon and, for this reason, the national indicator of GHG emissions, expressed in level terms, which we propose for our dashboard could be highly misleading if considered in isolation. Thus, in a dashboard it should always be complemented by some summary figures documenting total GHG emissions or, in lieu of complete data, CO_2 emissions. Our dashboard documents that the level of GHG emissions decreased in France and Germany between 2000 and 2008, whereas the level of worldwide CO_2 emissions increased substantially in the same period (Chart 4).

Obviously, an appropriate strategy limiting global anthropogenic GHG emissions requires a binding international agreement. Key elements of such an agreement should be a legally binding target for greenhouse gas emissions, an international emission trading system and an allocation mechanism that distributes emission allowances among the participating countries.

Even though quite different allocation mechanisms are conceivable, the principle of equality seems to be a good starting point for a fair distribution of the global budget. Therefore, equal per capita emission rights all over the world would arguably form a sensible basis for the allocation of national emissions budgets. Irrespective of its potential role in an allocation mechanism for globally traded emission permits, it would be valuable to inform policy makers and the general public about national **GHG emissions per capita**. Hence, we propose to include current GHG emissions per capita as a second GHG indicator in our dashboard. This indicator decreased in Germany between 2000 and 2008 from 12.5 tons to 11.7 tons. In France, the decrease amounted to approximately 10 % and the GHG emissions per capita came to 8.2 tons in 2008 (Chart 4).

36. The sustainability of (non-renewable) resources has been a hotly debated topic for decades among policy makers, scientists and the wider public alike. From the vantage point of economic theory, an emerging scarcity of non-renewable resources is primarily reflected in the evolution of their prices, and additional monitoring of physical measures does not seem necessary. But economic theory reaches beyond this hypothetical ideal, emphasizing the potential "over-use" of non-renewable natural resources that can occur as a consequence of **externalities** or of lacking **intergenerational fairness**. Therefore, beyond current prices we propose monitoring physical flows of non-renewable resource. This can be achieved by publishing indicators of non-renewable resource usage in production and consumption and their associated productivity, i.e. GDP relative to these measures. Our proposed first measure is direct material input (DMI) which comprises the total amount of raw non-renewable resources used in domestic production. Our proposed second measure is domestic material consumption (DMC) per capita, which measures the amount of resources consumed domestically by deducting exports from DMI. Prospectively, DMC should be enhanced to account for the resource content of imported and exported goods.

Applying both measures to France and Germany shows mixed results (Chart 4). Resource productivity increased steadily in France and Germany from 2000 until 2007. However, resource consumption per capita decreased in Germany while it remained fairly stable in France over the same time period. When taking into account the amount of resources embodied in imported and exported goods, Germany's resource consumption per capita decreased even more.

37. At least in a narrow sense, biodiversity is a form of capital that is required to produce services intended to satisfy human needs. Arguably, its preservation is essential for many desirable facets of current and future human existence, like food and nutrition security, medical progress or industrial raw materials. Moreover, ascertaining biodiversity is not only a global issue, but also relates to the stability of local ecosystems. Accordingly, due to its importance a **biodiversity indicator** should be added to our dashboard. Unfortunately, all existing indicators were developed outside the realm of economics, making it difficult to gauge whether they fully account for possible welfare trade-offs involved within and across generations. While we are not able to currently determine an explicit indicator that captures the economic dimension of biodiversity completely, we decided to include the bird index in our dashboard as the pre-

liminary fifth entry regarding environmental sustainability. This indicator decreased in France and in Germany between 2000 and 2007, suggesting that the variety of species decreased.

5. The road ahead

38. This **joint study** by the CAE and the GCEE thoroughly addresses a range of questions regarding economics and statistics which are as timely as they are fundamental. First, how can we improve our monitoring of **economic performance**? Second, how can we broaden our perspective from its current focus on economic performance to an assessment of **quality of life** more generally? And third, how can we design warning signals alerting us whenever the current manner of organizing our lives endangers **sustainability**? This study is not exclusively meant to be an academic investigation, venturing into the philosophical depths of assessing the state of mankind. Rather, it deliberately intends to represent a **pragmatic guide** to accounting for the current state of affairs.

Taking as its point of departure the SSFC Report, it discusses how comprehensiveness and accuracy might be optimally **traded off** against parsimony and cost to provide a reliable basis for regular, timely and digestible statistical reporting on human welfare. It is emphasized repeatedly throughout the report that a **desire for more information** not only implies higher cost of statistical reporting, but that the departure from the traditional measures of aggregate economic performance requires investing in the understanding of the **methodological intricacies** of statistical reporting. Typically, deeper insights can only be gained at the cost of greater complexity, and more sophisticated or theoretically satisfying concepts might not be used as directly as more primitive measures. In particular, while we might be able to construct new and insightful indicators for individual countries, they will often not be amenable to international comparisons, but only to comparisons over time within countries. In that sense, smarter statistics need smarter recipients.

39. The first and arguably most important conclusion of this study rests on a very similar consideration. We find ample reason to **dismiss** any single-indicator approach to measuring human progress as being insufficient. The study argues vehemently that life is too complex and the demands on statistical reporting are too diverse to allow a meaningful condensation of the current state of affairs into a **single comprehensive indicator**. While such a single indicator would emphasize parsimony and could be communicated easily, it could hardly do justice to the informational demands of modern democratic societies. Instead, we **suggest** that comprehensive statistical reporting should entail a compact **dashboard of indicators** (for a summary see the Chart 5 in the Appendix, page 27). Its basic idea is to provide a limited set of indicators that adequately cover all dimensions of human welfare which are relevant for short-term, medium-term and long-term policy decisions, and that is parsimonious enough to be used for the information of and communication to the wider public and policy makers alike.

Consequently, the dashboard we propose is rich enough to facilitate a meaningful discussion of the relevant facets of human welfare, but it is also not overwhelmingly extensive. Moreover, it provides a **balanced representation** of the three areas addressed by the key questions that inform our work. This approach acknowledges that monitoring material well-being is an

indispensable prerequisite for rational economic policy, that life is about more than material well-being, but that human progress in non-material aspects is quite difficult to capture, and that it is wise to take a long-term perspective by outlining the consequences of unmodified human behaviour.

Most indicators that we have chosen for our dashboard seek to encompass both current and future developments that impinge on present and future well-being. In particular, sustainability indicators are important as they convey a purely forward-looking perspective on developments and should be able to signal any potential need to take corrective action. Interdisciplinary discourse is urgently needed in the area of **environmental sustainability**, as the purely economic perspective on these issues is not comprehensive enough. A topic we paid particular attention to is **financial sustainability**. The unfolding of the recent financial crisis reminds us once more that there is a need to monitor certain aspects of financial and economic developments.

40. We consider our contribution as a **starting point** for further discussion and interdisciplinary discourse. This is even more important as certain dimensions are not limited to the field of social science and economics but inherently require a **multidisciplinary approach**. We therefore hope that our contribution fruitfully enhances the broad and lively debate about statistical reporting on the state of society, a debate that reaches far beyond the issue of material well-being. We strongly suggest bringing into this discourse the views of experts from the social sciences and from a wide range of other disciplines, elected officials and civil society.

As part of the envisioned public discourse, there should be a **regular review** of the appropriateness of the chosen indicator set. While care should be taken to avoid the list of indicators being subjected to frequent and politically motivated changes, an open debate about new challenges facing our societies and improved ways to capture progress would provide an important cross-check of whether current policies are consistent with emerging risks and opportunities. Finally, we suggest that governments present **regular reports** commenting on developments based on the dashboard. The confrontation between indicators of economic performance and current material well-being with indicators of quality of life and indicators of sustainability would bring the trade-offs facing policy-makers and society as a whole to the fore of the debate. Above all, this would help to overcome the problem that political decisions are often made with a very short-time horizon in mind.

Appendix

Chart 5

Dashboard for Monitoring Economic Performance, Quality of Life and Sustainability

Economic Performance (A)	Quality of Life (B)	Sustainability (C)
GDP per capita	Health: Potential years of life lost	Private sector net fixed capitel formation (% of GDP)
GDP per hours worked	Education: Students (ISCED 1-6) aged between 15 and 24 years	R&D investment (% of GDP)
Employment rate (15 - 64 age group)	Personal activities: Employees working on shift work	Cyclically adjusted fiscal balance (% of GDP)
Net national income per capita	Political voice and governance: Voice and Accountability	Fiscal sustainability gap S2
Final consumption expenditure per capita (including government consumption)	Social connections and relationships: Frequency of time spent with people at sport, culture, communal organization	Total private credit to GDP gap
Distribution measure of net income per consumption unit (income quintile share ratio (S80/S20); internationally harmonized)	Environmental conditions: Urban population exposure to air pollution by particulate matter	Real equity price gap
	Personal and economic insecurity: Not-at-risk-of-poverty rate	Real property price gap
		Level of greenhouse gas emissions
		Greenhouse gas emissions per capita
		Resource productivity (GDP relative to non-renewable Domestic Material Input, DMI)
		Resource consumption (non-renewable Domestic Material Consumption - DMC, per capita)
		Biodiversity (preliminary indicator: bird index)

References

Alesina, A., R. Di Tella and R. MacCulloch (2004) *Inequality and Happiness: are Europeans and Americans Different?*, Journal of Public Economics, 88 (9), 2009 - 2042.

Borio, C. and M. Drehmann (2009) *Assessing the Risk of Banking Crises – Revisited*, BIS Quarterly Review, March 2009, 29 - 46.

Costanza, R., M. Hart, S. Posner and J. Talberth (2009) *Beyond GDP: The Need for New Measures of Progress*, The Pardee Papers (4).

Easterlin, R. T. (1974) *Does Economic Growth Improve the Human Lot? Some Empirical Evidence*, in: David, P. A. and M. W. Reder (eds.): Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz, New York: Academic Press, 89 - 125.

European Commission (2009) *GDP and Beyond: Measuring Progress in a Changing World.*, Communication from the Commission to the Council and the European Parliament, COM(2009) 433 final.

Fleurbaey, M. (2009) *Beyond GDP: The Quest for a Measure of Social Welfare*, Journal of Economic Literature, 47 (4), 1029 - 1075.

Frey, B. S. (2008) *Happiness : A Revolution in Economics*, Cambridge, MA and London, England: The MIT Press.

Frey, B. S. and A. Stutzer (2002) *What Can Economists Learn from Happiness Research?*, Journal of Economic Literature, 40 (2), 402 - 435.

GESIS-ZUMA (2007) *German System of Social Indicators: Key indicators 1950 - 2005*.

Kuznets, S. (1934) *National Income, 1929-1932*, NBER Chapters, National Bureau of Economic Research, Inc., 1 - 12.

Layard, P. (2005) *Happiness: Lessons from a New Science*, New York: Penguin Press.

Nordhaus, W. D. and J. Tobin (1972) *Is Growth Obsolete?*, Economic Research: Retrospect and Prospect Vol 5: Economic Growth, NBER Chapters, National Bureau of Economic Research, Inc., 1 - 80.

Reinhart, C. M. and K. S. Rogoff (2010) *Growth in a Time of Debt*, American Economic Review, 100 (2), 573 - 578.

Schäfer, D. (2004) *Unbezahlte Arbeit und Bruttoinlandsprodukt 1992 und 2001*, Wirtschaft und Statistik (9), 960 - 978.

Schmidt, C. M. and S. C. Kassenböhmer (2010) *Beyond GDP Back: What is the Value-Added by Additional Components of Welfare Measurement?*, mimeo, RWI Essen.

Statistisches Bundesamt (2003) *Wo bleibt die Zeit? Die Zeitverwendung der Bevölkerung in Deutschland 2001/02*, Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt, GESIS-ZUMA and WZB (2008) *Datenreport 2008 : Ein Sozialbericht für die Bundesrepublik Deutschland*, Bonn: Bundeszentrale für politische Bildung.

Stiglitz, J. E., A. Sen and J. Fitoussi (2009) *Report by the Commission on the Measurement of Economic Performance and Social Progress*.

CHAPTER II

Economic Performance and Material Well-Being

Economic Performance and Material Well-Being

41. The assessment of well-being is a demanding and multifaceted endeavour. It involves, besides statistics, a variety of fields belonging to the realm of social sciences. Both the intensity of the public debate about this topic and the insights generated by research have increased steadily in recent decades, an evolution which is also reflected in official statistics. This chapter focuses on the aspect of **material well-being**, referring to the recommendations 1 to 5 of the SSFC report. More concretely, it carefully distinguishes two important themes, **economic performance** and material well-being. These two concepts are closely linked: economic performance **contributes** to material well-being in important ways, but is hardly the only factor affecting the current state of material well-being.

In its discussion of these matters, the present chapter seeks to strike a **balance** between complexity and parsimony. While ongoing progress has indeed provided professional statisticians and policy makers with much better tools, there is a risk that the general public could become progressively overwhelmed by ever more arcane and numerous statistics. To bridge the gap between producers and recipients, we aim at providing a **compact set** of indicators dedicated to economic performance and material well-being, respectively. This set could arguably form the basis for further discussion between the general public, policy makers and academics.

42. Monitoring economic performance and progress achieved in overall material well-being over time and across countries allows, in principle, the recipients of this information to address a range of **important questions**. The more ambitious these questions are, the more stringent the identification assumptions necessary for deriving the corresponding answer have to be. These challenges are:

- assessing a country's economic performance,
- assessing changes in current material well-being in a given country over time, and
- assessing the current level of material well-being in a given country or across countries.

Regular monitoring of economic **performance** is conceptually quite straightforward. Persisting problems to be addressed concern measurement and valuation issues, but their economic interpretation is undisputed. When it comes to assessing **progress** in current material well-being, which is defined as the well-being derived from the availability of goods and services, matters become more complicated since the link between performance and material well-being is intricate. Yet the state of the art in economics and statistics has evolved to a stage where meaningful statements regarding the changes in material well-being can be derived from the available data. It will be extremely difficult, though, to assess **levels** of well-being and compare them across countries as the corresponding measurement problems and the difficulties arising in the valuation of non-market goods tend to be overwhelming.

43. Our recommendations for advancing the state of affairs are similarly balanced. Ideally, to ensure continuous progress regarding both conceptual challenges (theory) and public policy design (application), **more resources** should be devoted to the collection of data on material well-being and to their analysis. Improving statistics is expensive, however, and official sta-

tistical systems typically operate under tight budgetary constraints. Consequently, we suggest that priority should be given to those items and issues that have a strong bearing on material well-being (**high marginal returns**) and which do not require major investment (**low marginal cost**), the so-called "low-hanging fruit". For other items, research should be encouraged both in academic institutions and government agencies, but here progress will be more of a long-term affair.

44. Chapter II is organized along the following lines. **Section 1** provides a broad overview of the challenges associated with the development of better measures of economic performance and current material well-being. **Section 2** maintains that economic performance is generally well captured by GDP statistics and demonstrates how the measurement of GDP can be improved. **Section 3** discusses labour market issues, which can be viewed as touching upon economic performance and material well-being alike. **Section 4** argues that documenting progress in terms of material well-being requires refocusing attention on a limited set of variables, most prominently on variables dedicated to distributional issues. **Section 5** concludes.

1. Economic performance and current material well-being

45. As **GDP** is a measure of a country's overall production for any given year, it is a **reliable**, albeit still **imperfect gauge** of that country's **economic performance**. This is the justification for the great attention which both the general public and policy makers pay in all advanced economies to the regularly published GDP figures. Yet GDP in particular and national accounts in general are dogged by certain well-known deficiencies related to the measurement of economic activities.

Moreover, as Chapter I already discussed concisely, the measurement of GDP does not address all aspects which are relevant for the **material well-being** of an economy. While the general public and many policymakers unduly regard GDP as a measure of material well-being, this interpretation ignores the indisputable fact that production is not the ultimate goal of a society. Therefore, production-based measures need to **be complemented** by a broader set of indicators if the aim of the exercise is to assess well-being.

Measurement problems

46. Some problems associated with measuring GDP are well known. Various **non-market outputs**, such as household activities and services provided free of charge, are systematically overlooked. The **underground economy** is difficult to capture, particularly certain criminal activities like drug trafficking, although several attempts have been made to harmonize the coverage of the underground economy at EU level in order to obtain comparable GDP measures for EU budget and deficit purposes. Some elements of GDP are **fragile estimates**, particularly those of the volume of publicly provided services and of the quality incorporated into products. Finally, some expenditures are unequivocally counted as positive contributors to economic performance, while the **negative externalities** associated with them – such as environmental damage – are neglected. These problems have to be alleviated if the assessment of current economic performance is to be improved.

47. GDP excludes activities that are not provided through the market, such as **household production** (childcare, housekeeping, preparing meals, and care for the elderly) and voluntary or unpaid services. As a result, GDP **understates output**. Calculations based on information about time spent on these activities and a valuation of working time at the standard cost of a paid housekeeper even suggest an upward adjustment of, e.g., French or German GDP of about one third. This lack of comprehensiveness is all the more problematic as market boundaries have greatly expanded over the years. Many services formerly provided by family members are now purchased on the market. The result is an increase in output and income – giving the impression that household living standards are rising, although in fact the fundamental circumstances might not have changed: all that happened was a shift from services produced by households to services produced on the market.

48. Current assessments of **services** are not satisfactory. This holds in particular for government services provided in kind, such as healthcare and education. Deriving appropriate **market prices** for **outputs** linked with a healthier or better educated population are scarcely available. Similar problems arise in the context of estimating the contribution of financial intermediary services (FISIM). Therefore, statisticians normally resort to prices of **inputs** such as earnings of doctors, nurses, and teachers. But, among other problems, this methodology ignores quality improvements in public services. This weakness is all the more problematic given the substantial weight of public services in GDP (in 2009: 18 % in France and 19.6 % in Germany), a share that is trending up steadily in today's economies. Most importantly, these difficulties prevent the meaningful construction of **international comparisons**. For example, if a country has opted for public provision of most of its healthcare services, and if these are underestimated by the valuation method described, the country will seem to be less rich than a country in which the same services are provided by the market and valued at market prices.

49. Official GDP estimates tend to omit important parts of the **underground economy**. For instance, transactions such as illegal drug trade remain unreported, causing GDP to be underestimated. Apparently all attempts to incorporate the valuation of the informal economy, which, one way or another, have had to rely on indirect estimates, have led to **major revisions** of the official GDP figures. An extreme example is provided by Colombia, which revised its GDP upward by 16.5 % in 1994, in particular by including an estimate of the production of illegal crops. Consequently, ignoring factors such as this may undermine international comparisons of economic performance. Even so, due to the inherently tremendous **measurement problems**, GDP figures corrected for illegal activities like the drug economy should always be treated with considerable caution.

50. In addition, **quality improvements** and supply of **new products** are difficult to account for, possibly causing GDP to understate true economic growth. The chain of causation is clear, since underestimating qualitative improvements means **overestimating prices** and hence **understating real income**. Providing the necessary adjustments to GDP is a delicate matter, however. Most importantly, one needs to distinguish between new models and varie-

ties of previously existing products on the one hand and genuinely new innovative products on the other.

In the European definition of the harmonized index of consumer prices (HICPs), **new varieties of products** are introduced as a replacement and the prices are adjusted accordingly. Past experience suggests that such quality adjustments tend to have a perceptible impact on the assessment of real activity. In some countries, especially in the information technology sector in the 1990ies, the growth in "production" has resulted more from qualitative improvements in the products produced and consumed than from an expansion of their quantities. Quality improvements of the goods traced in consumer price statistics are still perceived and treated in different ways in different EU countries. And the difference may not average out across the goods and services covered by the indices: On the EU-level, they are likely to cumulate to differences well in excess of 0.1 percentage points.

New innovative products are introduced into the HICPs by addition as soon as they are relevant for consumers. The price of the new product is collected in addition to the products already observed and the weights for the relevant consumption category are adjusted accordingly.

51. Finally, if one takes the view that, besides serving as a measure of economic performance, GDP is also intended to be a measure of material well-being, it should certainly not capture expenditures that would usually be associated with a decline rather than an increase in current material well-being. Security expenditures are an example of such "**defensive costs**," to use the term adopted by Nordhaus and Tobin (1973). That is also the case with medical and repair bills for automobile accidents, commuting costs, and household expenditures on pollution control devices such as water filters. The SSFC Report suggests that these expenditure categories should be treated as investments or intermediate consumption rather than consumption expenditures. Most importantly, GDP ignores non-compensated externalities such as **damage** to the environment resulting from the depletion of natural assets and the production process. By failing to account for the negative effects associated with higher production – such as more pollution – GDP not only completely ignores sustainability issues but also tends to **overstate** current material well-being.

52. An additional challenge to the appropriate measurement of GDP is posed by continuing **European integration**. GDP explicitly measures economic activity within a single country, which in previous decades was demarcated by customs frontiers and the sovereign realm of its own currency. In today's intra-EU trade there is **no physical registration** of imports and exports, due to the abolition of intra-EU customs barriers and the introduction of the euro. Exports and imports are solely recorded statistically and matched with national turnover statistics. Consequently, in the Intrastat-System considerable differences might occur, as, for example, the recorded imports of a member country could deviate from the recorded exports of all other member states to that country. This can, at least in the short run, distort GDP measures, exacerbating the challenge of monitoring and fiscal policy coordination of the member states.

From production to material well-being

53. Apart from these well-known deficiencies in capturing economic performance, production-based measures such as GDP fail to capture some important aspects of material well-being. Measures based on **income** or **consumption** are arguably more suitable than GDP for approximating material well-being. Moreover, aggregates that are closer to **households** than is GDP can provide valuable additional insights into growth patterns and developments of material well-being. Looking, for instance, at France, Germany and the EU-27 in the 2000-2009 period, different variables tell different growth stories (Table 2). While all measures shown here indicate a higher growth rate for France and the EU-27, the major difference is the stronger French performance for household disposable income and final consumption.

Table 2

Growth in France and Germany measured by alternative indicators

Average annual growth rates between 1999 and 2009 (%)

Indicators	France	Germany	EU 27
Gross domestic product per capita	2.7	1.8	2.8
Gross domestic product per hours worked[1)]	3.3	2.4	3.2
Gross national income per capita	2.6	2.0	2.8
Net national income per capita	2.4	2.0	.
Private final consumption expenditure per capita[2)]	3.2	1.9	2.8
Net household disposable income per capita[3)]	3.3	2.0	2.9

1) Between 2000 and 2008.– 2) Private households and non-profit institutions serving households.– 3) Including non-profit institutions serving households.

Source: EU

Data

54. As the **value of leisure** is not included in GDP calculations, international differences in GDP or GDP per capita or GDP per hours worked might at least partly reflect disparate preferences for goods and leisure. Thus, an international comparison of levels is fundamentally problematic. This is a different matter, however, when the aim is to compare economic performance or also changes over time in material well-being. In this case, since it may be justifiably assumed that **preferences** change relatively slowly, it makes good sense to undertake a direct comparison of progress made in material well-being. As the greatest conceptual difficulties arise when attempting to compare levels of material well-being, this is the research area where further improvements are most desirable. Due to the nature of the comparability problems, this progress must be achieved outside the national-accounting framework, via **time-use surveys** and supplementary satellite accounts.

Another important limitation of highly aggregated figures is that they disregard **income disparities** between high and low- income households, between domestic and foreign owners of production factors, and between workers and domestic owners of capital. There is no doubt that different distributions of the income created by the production process can lead to different degrees of well-being. In particular, where there is a tradeoff between equity and economic performance, one would need to gauge the value a society assigns to these conflicting goals. This calls for separate treatment of economic performance and changes in material well-being.

55. Material **wealth** plays a dual role in the determination of material well-being. First, changes in net wealth can signal a deterioration or an improvement with respect to future availability of goods and services. An important lesson of the current financial crisis is that standard measures of economic performance and current material well-being can display high growth rates that largely camouflage unsustainable increases in indebtedness relative to income and wealth. Looking at wealth from this perspective thus raises the issue of **sustainability**, which we shall accordingly discuss in Chapter IV. Second, the current level and distribution of income and wealth is a factor in determining current levels of material well-being, as will be discussed later.

56. From this discussion, it follows that **two strategies** for improving the current set of indicators have to be implemented simultaneously. Clearly, the existing deficiencies of GDP as a measure of economic performance are not severe enough to discard GDP and the measures derived from it altogether. Rather, the first approach should be to retain these measures and to improve upon their current state through **appropriate adjustments**. In implementing this strategy, one has to decide the order in which the deficiencies should be addressed. Defining priorities will require balancing information on the magnitudes of these problems and the cost of achieving noticeable improvements. Second, it has become apparent that measuring changes in the current level of material well-being requires a broader set of indicators than GDP alone, indicators which capture discrepancies between consumption, income and production measures as well as distributional issues (Box 1).

Box 1

How to capture distributional issues in the national accounts: breaking down the household account by household category

Neither GDP nor other national-accounting aggregates can capture changes in resource distribution, and they do not allow a breakdown of the changes by household category. To supplement the macroeconomic analysis based on the national accounts accordingly, one needs microeconomic data gathered from household surveys. Most importantly, these provide sufficiently rich information for constructing inequality indicators for each household category. But the differences in definitions and methods can also create divergences between macroeconomic and microeconomic data, thereby "scrambling" the messages sent by the various indicators. Improving consistency between these two sources should therefore be a priority for each statistical system, but it is an arduous undertaking and achieving perceptible progress has to be perceived as a medium-term goal.

Along these lines, statistical offices throughout Europe seek to supplement regular national accounting statistics with appropriate data that reflect the distributional perspective of national income. Both INSEE in France and the Federal Statistical Office of Germany (Destatis) are pursuing projects to establish personal income distribution data which facilitate a detailed insight into the situation of various types of households. For example, INSEE has already published data on the breakdown of the 2003 household account (income, consumption, and saving) by household category (INSEE, 2009). As a consequence, disposable income, consumption expenditures and the saving ratio for different household categories are available for the year 2003. These data, coherent with national account data, are provided according to a decomposition by quintile of disposable income by consumption unit, household composition plus age and socio-occupational

category of head of household. In further work, INSEE also plans to publish a breakdown of the household balance sheet by household category and ten-year trends of the household income by household category for the period 1997-2007 (Appendix, pages 56 f.).

2. Making GDP a better measure of economic performance

57. Economic policy makers unquestionably need an economic-performance indicator for **short-term decision-making**. Macroeconomic policy frequently operates with a time horizon of one to two years, and from this perspective GDP, as an indicator of current value added, is arguably a most informative gauge of economic performance. In this context, all the measurement problems raised by the SSFC Report are only of moderate importance. Of course, even in this area of economic policy, one typically goes "**beyond GDP**" by analyzing data on unemployment, inflation, short-term business activity and consumer or business sentiment. And although the usefulness of GDP is limited from a medium-term perspective, it still remains a viable indicator of **medium-term performance**. Thus, in conceptual terms, GDP remains the cornerstone of economic-performance assessments. Nevertheless, it should be improved in various directions.

However, not all the issues discussed in the previous section can be addressed with similar intensity and at the same time. From our point of view, the most promising starting points for improvements are (i) improving the measurement of **service output** in general, and of **government services** in particular, and (ii) making progress in **measuring quality** improvements. These are the issues we discuss in the following paragraphs. Specifically, a work programme should be defined to determine output measures for public services. In our assessment, other problems of GDP measurement are secondary in nature, and any effort to correct for them might entail insufficient value added or even lead to a loss of reliability.

Services

58. In today's economies, services account for as much as two-thirds of total production and employment. Given their economic weight, it is unfortunate that we have such an imperfect grasp of both the precise **volume** and **quality** of services, since these two components play an important role in calculating GDP. Yet while prices and volumes of services are often more difficult to measure than those of goods, the human resources – in terms of statisticians – available for tracking the service sectors remain relatively modest. Among the areas where appropriate **output price indices** are desperately needed are business services such as financial intermediation, health services, and research and development.

But perhaps the most promising avenue concerns the improvement of measuring **in-kind services**. In France and Germany, these reached € 391 billion and € 516 billion in value added in 2009, respectively, with health, education and welfare services accounting in both countries for around 30 % of total employment (Table 3). Traditionally, the output of non-market services provided by governments is valued in **nominal terms**, by summing the expenditures incurred in supplying the service. These expenditures pertain to the cost of labour, intermediate consumption, fixed-capital consumption, and taxes linked to the production of the ser-

vices. Yet if we want to interpret their sum from a standard-of-living perspective, we would have to assume that costs equal the **valuation** of the service by its recipients. But how can we be sure that this is the case, as no actual market transactions are observed? This is an intellectual puzzle that recent research in statistics has addressed quite successfully, however.

Table 3

Value added and employment by economic sector in France and Germany (2009)

Economic sector	France				Germany			
	value added		employment		value added		employment	
	bln Euro	%	1,000 persons	%	bln Euro	%	1,000 persons	%
Agriculture	30.0	*1.7*	793	*3.1*	17.3	*0.8*	866	*2.2*
Industry	213.4	*12.4*	3,254	*12.7*	474.4	*22.2*	7,814	*19.4*
Construction	111.0	*6.4*	1,787	*7.0*	92.1	*4.3*	2,200	*5.5*
Services	1,367.4	*79.4*	19,726	*77.2*	1,556.8	*72.7*	29,385	*73.0*
Thereof:								
Market services	976.1	*56.7*	12,140	*47.5*	1,040.3	*48.6*	17,004	*42.2*
Non market services	391.3	*22.7*	7,587	*29.7*	516.4	*24.1*	12,381	*30.7*
Thereof:								
Education	93.8	*5.4*	...	...	100.3	*4.7*	...	...
Health	101.3	*5.9*	...	...	171.8	*8.0*	...	...
Social welfare	56.3	*3.3*	...	...	113.4	*5.3*	...	...
Administration	139.9	*8.1*	...	...	130.9	*6.1*	...	...
Total	**1,721.7**	***100***	**25,561**	***100***	**2,140.6**	***100***	**40,265**	***100***

Sources: Destatis, INSEE

Data

59. Until the base-year 1995 national accounts, the total value of the services provided was approximated by taking the total value of the factors used to produce them: this was known as the **input method**. The change in prices was therefore taken as equaling the change in the cost of the production factors, a choice that precluded productivity gains. An EU Regulation of December 2002, implemented in 2006, invited EU countries to use an **output method** for non-market education and health services, based on direct indicators of the volume of the service produced. To apply the method, one needs to collect direct indicators of production volume at the most detailed level possible, so as to calculate elementary indices that will then be weighted by the production costs estimated at that level of detail.

However, the adoption of such a method and the specific choices regarding its implementation, in particular the assessment of **quality effects**, are still being debated at the international level: Merely describing the output of non-market services at a detailed level is does not suffice to track the quality of services provided. The final outcome in the education sector is the increase in the skills of students trained in schools and universities. The final outcome in the medical sector is better health achieved through medical care. But these outcomes are difficult to observe because the results obtained in education and health depend on other factors such as the pupil or student's cultural environment (Cutler et al., 2006) or on the patient's life style.

The qualitative change in the supply of these services should therefore be measured by the **marginal contribution** of education or health to the population's level of knowledge or

health, controlling for all the other influences. These methods are complex to put into practice. They require gathering additional information on individuals and their environment through surveys, and performing **econometric estimations** to determine the effects. This kind of exercise falls more into the category of original research work rather than that of routine national accounting. The statistical offices of the United Kingdom and Italy have started to investigate this field, but this is still work in progress.

60. Using measures of outputs that capture the number of patients treated or the number of students having received training has **major effects** on **reported GDP**. According to INSEE, the French economy grew at an average 2 % a year between 2000 and 2006 if the output method is used to calculate the volume of non-market health and education services, versus 2.15 % using the input method. The UK economy expanded by an annual average of 2.75 % between 1995 and 2003 on an output basis, as against 3 % if the country had continued to use the input method (Atkinson, 2005). The switch from the input method to the output method produces slight downward adjustments for France and the United Kingdom, but an alteration upwards for Denmark.

61. These results show the challenge involved in seeking a better measurement of **non-market services** provided by **government**. Practices should be harmonized at international level to obtain comparable results. In our view, the output method without quality effects – which consists in a detailed description of service production – is a better choice than the input method for measuring the volumes of education and health services (Box 2). This choice would significantly improve the measurement of GDP. The close tracking of health expenditures by governments in order to control them better should supply, as a by-product, the administrative sources necessary for a reliable valuation of those services. By contrast, we believe that the assessment of quality effects for health or education is more suitable as a research topic and should be addressed only in satellite accounts. We also recommend that the satellite accounts should provide a parallel estimate of the output of education and health services using the input method so as to have some measures of productivity growth in these services.

In theory, the output method with quality effects is also relevant for other government services provided in kind, such as social work, recreational facilities and activities, and security. But it is far more difficult to implement owing to a lack of information in these fields. We therefore recommend starting working on education and health services.

62. The **depreciation** of fixed capital plays a major role for non-market producers such as the government sectors, because of the additive cost-accounting approach. In most cases, depreciation needs to be estimated by models, because it is not recorded on a micro-basis (public fiscal or “cameralistic” [single-entry] accounting) or because business-accounting concepts differ from national-accounting concepts. We see a need to harmonize practical methods used to calculate depreciation for collective non-market services. More generally, there is a need for a more satisfactory evaluation of services provided by investment. For instance, US GDP already includes depreciation of military equipment.

A more serious problem is that conceptual differences distort international comparisons. Military consumption is still considered intermediary consumption in the French, German and European national accounts according to ESA 1995, whereas it is included in fixed capital formation in US national accounts. This investment produces defense services in accordance with the SNA 2008.

Box 2

Evaluation of individual non-market education and health services in France

France adopted the output method for non-market education and health services in its base-year-2000 national accounts. For this purpose, INSEE gathers direct indicators of output volume at the most detailed level possible. INSEE can then compute primary indices and weigh them by costs estimated at that level. More specifically, INSEE estimates the volume of non-market education as the number of teaching hours by grade and subject multiplied by the number of students. The costs are those of education for general government, published by level and programme in the satellite account for education.

For healthcare, INSEE calculates an output volume index by weighting by their relative costs the various indices for several hospital-activity indicators supplied by the annual statistical survey of healthcare facilities (Statistique Annuelle des Établissements: SAE) as well as activity indicators obtained through the "programme for medicalization of information systems" (Programme de Médicalisation des Systèmes d'Information: PMSI). More specifically, the output of non-market healthcare by hospitals is broken down into a set of treatments, which INSEE can aggregate into three broad categories: short-stay care or medicine-surgery-obstetrics (MCO), medium-stay care, and psychiatry.

For short-stay care (MCO), the method uses PMSI data at detailed level, which supplies the number of stays (for which quantity indices are calculated) and unit costs per stay (used for weighting) for 600 "Diagnosis-related groups" [DRGs] (in French: Groupes Homogènes de Malades [GHM]). This facilitates constructing a true volume index. For follow-up and rehabilitation care (medium-stay care) as well as psychiatry, in the absence of fuller information, INSEE simply takes quantity indicators (number of admissions for full-time and part-time hospital care) drawn from the SAE. In the area of education Destatis applies a similar procedure to that of INSEE. For health services Destatis uses the DRGs to determine a price index for deflating the nominal values.

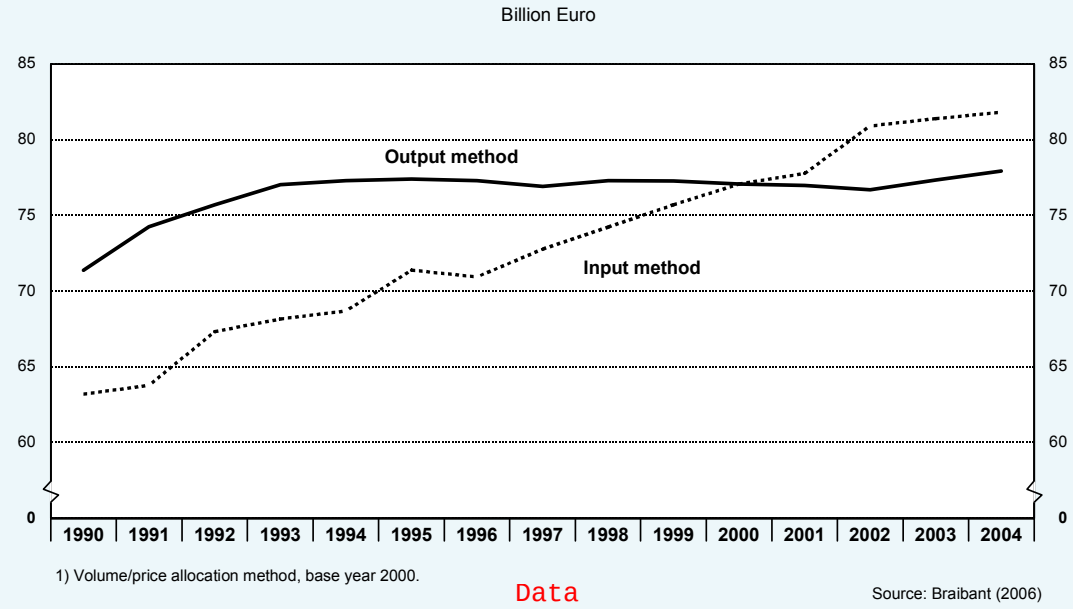
This procedure provides a measure of the implicit quality change in the production of education and health services, but does not yield an evaluation of the quality of education provided. There are drawbacks to this approach. For public education, the results have therefore been adjusted by an education quality indicator, which reflects the annual number of pupils or students who successfully complete each course level.

The new assessments using an output method have led to a downward revision of the volume change in the production of non-market education in recent years. The positive real change in the base-year 1995 accounts reflected the increase in resources applied, notably the improvement in teacher skills and the increased number of teachers. In the base-year 2000 accounts, the stagnation in volume terms observed since 1996 is linked to demographic changes (decline

in number of pupils/students enrolled in some programs) not offset by an increase in the exam pass rate or by a more frequent move to a higher grade (Chart 6).

Chart 6

Education services by different statistical methods for France[1)]



1) Volume/price allocation method, base year 2000.

Data

Source: Braibant (2006)

63. Bid-ask spreads, which are the interest margins observed in securities management activities, are not taken into account in measuring **financial intermediary services** (Financial Intermediation Services indirectly Measured: FISIM), and this represents another deficiency. From a national-accounting perspective, measuring activity is trickier in banking than in any other sector because many services provided to customers are not explicitly charged for. Bank revenues are measured by the value of services explicitly charged for through fees paid by financial and non-financial customers, but also by the value of **implicit services** rendered by banks and effectively paid for by customers. This is especially true of financial intermediation services connected with managing customer deposits and loans. These services are mainly remunerated through the interest margin that banks earn by lending at a higher rate than the one they pay to obtain funding. The margins are measured in accordance with European regulations, which set the procedures for calculating financial intermediation services indirectly measured.

But traditional bank financing, based on managing deposits and lending to customers, has given way to interbank transactions and securities market intermediation. The bid-ask spreads should be classified as services implicitly charged for, and hence part of FISIM. They may generate substantial gains because of the volumes traded, although the bid-ask spreads for market-makers are relatively narrow (€ 10.0 billion in 2008). In practice, up to now, they have been treated as capital losses or gains linked to changes in the market value of securities. To improve the estimation of the value added of the financial sector and hence of GDP, one could

recommend including bid-ask spreads in FISIM despite the major measurement problems they pose.

Quality changes and trade issues

64. The growing share of services in the total economy, combined with the production of ever more complex goods, makes it increasingly difficult to measure output volumes and, consequently, economic performance. Today, the **quality** of many products is complex, multi-dimensional, and subject to rapid change. This is clearly the case for items such as cars, computers, washing machines and financial services. Tracking qualitative change is therefore a formidable challenge, but essential to measuring real GDP, income, and consumption, which are decisive factors in economic performance and people's material well-being. These adjustments are delicate, given their significant consequences, and they involve substantial work.

On the EU-level, for the HICPs, quality change is said to occur whenever the change in specification has resulted in a significant difference in **utility** to the consumer. Quality adjustment then means to increase or reduce the observed price difference by a factor or an amount equivalent to the value of that quality change. Quality adjustments in the HICPs should be based on **explicit estimates** of the value of the quality change. Where no estimates are available, price changes should be estimated as the entire difference between the price of the substitute and that of the item it has replaced. EU Member States are required by legal standards to avoid the so-called **automatic linking** method, which is equivalent to the assumption that the differences in prices between two successive models are wholly attributable to quality differences.

Despite these existing legal standards, differences between HICPs may arise because the same changes in the physical characteristics of an item are still perceived and treated in different ways in different countries. Eurostat and the EU Member States are currently involved in developing and rating quality adjustment methods. So far, **standards** have been agreed for clothing, footwear, books, recorded media, computer games, and for cars and other vehicles.

65. It is essential to ensure that quality effects are properly measured in order to obtain an accurate assessment of GDP. But we must also be pragmatic. What counts is not so much having sophisticated methods, but rather having **robust methods** which are harmonized across countries to ensure the comparability of results. Case-by-case studies on problematic products, supervised by international entities with the aim of identifying best practices, therefore seem promising. One should also be careful not to introduce excessive sophistication into methods that are relevant in research work but are harder to implement in statistical production work or by countries with less advanced statistical systems.

These methods should be reserved for measuring prices of goods or services that account for large shares of consumer budgets (such as automobiles) or capital formation. The calculation of global productivity depends heavily on analyzing the changes due to price movements and volume movements, the latter incorporating quality effects. As many such goods are traded

worldwide, cooperation among statistical offices could make the methodological investment less costly.

66. Another shortcoming is that intra-EU imports and exports are not recorded below a certain threshold. In **intra-EU trade**, there is no registration of imports and exports below a certain value (Box 3) owing to the abolition of intra-EU customs barriers. The missing flows need to be estimated in order to prepare national accounts.

Box 3

Measurement of European Union trade

In January 1993, following the establishment of the Single Market, the European Union (EU) introduced a new system called Intrastat for collecting statistics on trade between member states. Based on declarations by firms of transactions above a certain threshold value, the new harmonized community system affects EU merchandise trade statistics in many ways:

- Comparability with pre-1993 figures has been lost.
- Coverage under the new system is less exhaustive than that of the customs-based system.
- Unlike the previous system, intra-EU imports are about 5 % below intra-EU exports. Possible causes of this asymmetry problem include differences among member states regarding adjustments for non-response, confidentiality, triangular trade, and thresholds.
- The variation in thresholds illustrates the problem. Prompted by the desire to reduce the response burden on businesses, thresholds for intra-EU trade range from about € 30,000 to over € 600,000; the recommended threshold for extra-EU trade is only € 800.

At present, there are pressures to raise thresholds even further so as to reduce the data collection burden on respondents.

Deficiencies that are harder to correct

67. The discussion of topics above highlighted areas of promising statistical research and possible progress towards making GDP a more accurate and more reliable construct for capturing economic activity within a country. These improvements can be achieved in a short to medium-term perspective with reasonable effort. By contrast, correcting other deficiencies in measuring GDP might impose **prohibitive costs**, which leads us to conclude that they should not be addressed with high priority. For instance, we question the likely success of efforts intended to capture the production of those **non-market services** which households provide to themselves. While conceptual purity is always to be commended, the inclusion of such services would pose severe measurement problems. Most importantly, given the large amounts involved – for France, for instance, they are estimated at approximately one third of GDP between 1995 and 2006 – the adjusted GDP figure might be severely distorted if the attempt at accurate measurement failed.

In fact, in practical applications the figures obtained vary considerably depending on whether one decides to **value** unpaid work performed in the household using (i) the wages of household employees (specialized or non-specialized) or (ii) the wages that the persons actually

doing the work earn or could earn in the job market (opportunity cost). Moreover, the real change in such services tends to be assessed very differently depending on whether or not one also considers possible variations in **productivity**. Furthermore, the calculations should ideally be based on internationally comparable time series on time use in different countries. While such work is in progress in the United States and several European countries, it is still non-existent in many other countries.

Moreover, the adjustment would tell us nothing more about short-term economic developments, as the changes in the output of such services are practically impossible to track from year to year. By contrast, capturing such services would be wholly justifiable in a **satellite account**. In Germany, satellite accounts for household production have been constructed for 1991 and 2001.This additional information would be valuable in analyzing long-term changes in a single country or for international comparisons in connection with assessing living standards.

68. Similarly, it does not appear very fruitful ultimately to attempt to adjust GDP for the estimated value of **defensive costs**, as was proposed by Tobin and Nordhaus in the 1970s. First of all, the concept itself is hard to pin down. In principle, these costs would consist of all goods and services that ought to be subtracted from output because they do not directly promote personal well-being. Examples include expenditures relating to road accidents, prisons, and oil-spill clean-ups. But could we not argue that healthcare and automobile repair expenditures do contribute to the well-being of society if we accept that accidents are inevitable with cars on the road? Likewise, if we accept that, unfortunately, crimes and felonies are inevitable in a society; prisons contribute to citizens' peaceful existence and well-being.

A decision to subtract these expenditures from output may undermine the principles of national accounting, which does not rest on ethical judgment. Hence, for instance, the output of the underground economy should, theoretically, be included in GDP. Removing defensive costs from the determination of GDP would be all the more regrettable as we can approximate material well-being by other national-accounting indicators such as household income or consumption. Such household-centered indicators do not include defensive costs, which consist mainly of "collective" expenditures by general government not directly related to households.

The issue is even more delicate when the defensive costs merely serve to remedy the earlier or concomitant deterioration of **economic stocks** or economic natural assets, such as damage caused by pollution. In either case, however, the outcome should not be a decrease in GDP. For example, when an earthquake destroys buildings, the estimated value of the destruction should be subtracted from the value of the building stock. Later on, the value of the reconstructed buildings will increase the stock. In this way, neither GDP nor GNP will be affected, since we are not dealing with a flow linked to production or consumption, but the building-stock value will have been effectively adjusted. At the time of (re)construction, the construction industry's output will have as its counterpart a gross fixed capital formation in buildings and a rise in GDP. In fact, this example is an argument not against the method for determining GDP, but rather in favour of the preparation of expanded balance sheets.

69. The deterioration of **natural assets** (atmosphere, sea) cannot be treated in the same way, since these assets are not included in the national-accounting balance sheets. Natural-asset degradation by an economy is a consumption of natural assets. This is an area of statistical accounting where extremely large valuation problems pose quite prohibitive obstacles.

Intermediate conclusions

70. In short, our analysis of the various pitfalls and shortcomings associated with GDP measurement leads us to the following conclusions. To start with, there is no obvious need to turn GDP from a powerful indicator of **economic performance** into an indicator of material well-being or growth sustainability. To deal with these issues, better suited indicators already exist within the national accounting framework and elsewhere. While indicators of material well-being are the topic of the next section, sustainability will be covered in chapter IV.

By contrast, attempting to measure output without stepping out of the national-accounting framework entails **genuine problems**. Two prominent examples are the unsatisfactory measurement of the volume of publicly provided services and the need to adequately incorporate quality improvements in products and services. Other issues of a similar nature are the imperfect coverage of certain outputs that should, in theory, appear in the national accounts, such as the output of domestic services by households or the underground economy. In principle, all of these weaknesses should be remedied.

71. These are the deficiencies that statisticians will need to correct in the years ahead. In setting their **priorities**, statisticians should concentrate their current efforts on the most attainable objectives (the "low-hanging fruit"), meaning those that will provide the largest **marginal information gains**. It seems less promising, for instance, to dedicate major resources to capturing the underground economy in detail since the estimates can only be made through statistical imputations and the results will always be fragile. By contrast, we can expect far greater gains from a correct measurement of publicly provided services. The same holds true for a better volume-price breakdown that takes quality effects into account. As an overarching principle, the quest for an exhaustive valuation of all output components runs the risk of raising that share of GDP which is evaluated by imputation, thereby making the results less **robust**. That is why the measurement of domestic services might be more appropriate in a satellite account than in the central framework.

Improving GDP along these lines would be a highly valuable – indeed essential – component of a dashboard of indicators seeking to cover aspects of economic performance. Whenever demographic changes are an important phenomenon or where an international comparison of economic performances is desired, there is a need to adjust the reported GDP figures for the size of the economies under scrutiny. We therefore advocate always presenting the growth rate of **GDP per capita** in the dashboard. To add a measure of productivity as a major source of economic performance, one should also consider reporting the growth rate of **GDP per hour worked**. These two indicators appear to provide the best trade-off between a fair indication of a country's overall economic performance and methodological robustness. This conclusion already holds for presentation according to current standards of measurement. But

naturally, the two indicators will provide even better information once they take account of the main flaws discussed here.

3. Labour market issues

72. As argued in the previous section, we maintain that GDP is a powerful indicator of a country's economic performance. As a measure of total output, it comprises the total amount of goods and services produced within a certain period, with capital and labour as the principal input factors. And yet, labour is more than simply a **factor of production**. It seems fair to argue that in all countries, almost all working-age persons want to have a job, not only in order to gain access to consumption, housing, and social-insurance benefits, but also to retain a decent social status. Most importantly, a high probability of employment seems to be a **major prerequisite** for ascertaining material well-being. Thus, employment and unemployment arguably occupy somewhat of a hybrid position between being elements of economic performance and being facets of material well-being. This is reason enough for us to treat them in a separate section of this chapter.

73. Several indicators are conceivable for capturing the state of affairs on the labour market. The first indicator that springs to mind is the **unemployment rate**. However, this is hardly the most sensible gauge for our purposes, since it is heavily influenced by country-specific legislation and programmes to combat joblessness. Also, whenever unemployment is too high and long-lasting, workers might quit the labour market, making inter-country comparisons particularly unreliable. We therefore propose using a more direct indicator in our dashboard, namely the probability of being employed at working age.

More concretely, we suggest using the **employment rate** in the population aged **15-64 years**. This basic indicator has already gained widespread acceptance in labour economics and statistics. Admittedly, the age limits chosen are debatable. In highly developed countries, a large proportion of youngsters above 20 years of age are still studying. At the other end of the age range, the retirement age has effectively declined below 65 in recent years, but is now rising again, among other influences because of improved life expectancy. For the time being, however, the employment rate for the age range 15-64 still seems to be the best indicator. At some point in the future, this indicator should be harmonized according to the Europe 2020 strategy, which refers to the age group 20-64.

While such an indicator admittedly does not tell us anything about job quality or whether jobs match people's expectations, it nevertheless means a lot when looking for a job or being exhausted by long periods of job search. Of course, there can be other interpretations of this indicator, as it is a powerful quality-of-life driver. It reflects, to a large extent, the choice between private time and work. It is also a sustainability indicator as it is an important parameter for the long-term future of retirement plans and public finances.

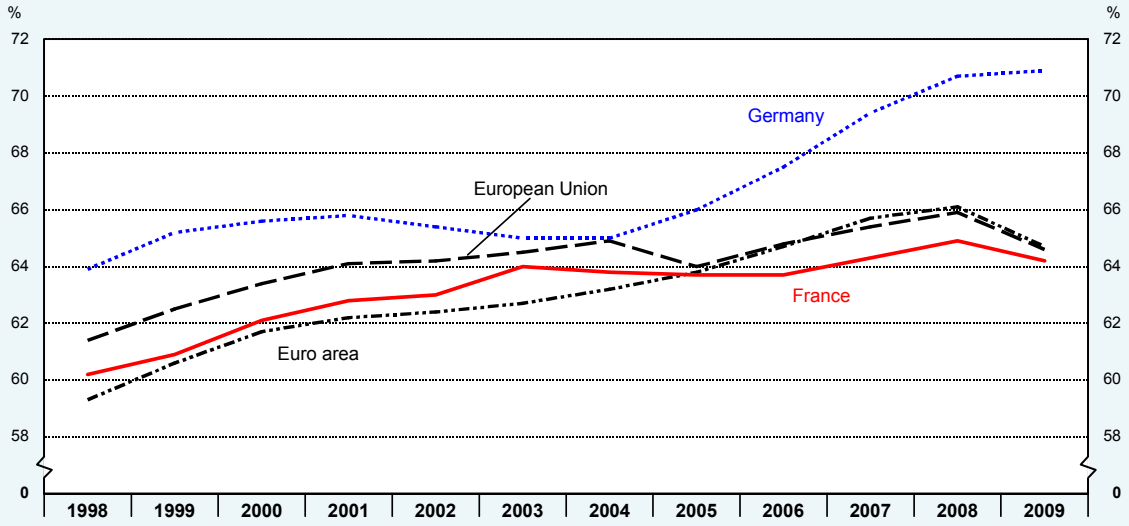
74. The 15-64 employment rate experienced over the 2004-2008 period displays a mild rise in France and in the EU and a strong increase in Germany (Chart 7). In 2009, in the context of the global crisis, employment fell moderately in France and more significantly in the EU,

while it actually rose somewhat in Germany. At the current juncture, the employment rate in Germany reaches 71 % of the population aged 15 to 64 years, compared with 64 % in France. The employment gap between the two countries has therefore markedly widened over the past few years.

Chart 7

Employment rate in Europe

Employees in relation to population aged between 15 to 64 years



Data

Source: EU

4. Defining a wider set of indicators for material well-being

75. We now turn to the task of measuring changes in material well-being. **Three dimensions** will be successively considered: income, consumption, and wealth. In our analysis, we shall bear in mind that while significant differences in preferences persist across countries, they are of a structural nature and thus tend to be relatively permanent. As a result, although measuring the level of well-being is an intricate matter, it is highly unlikely that comparability problems interfere significantly with the measurement of changes in material well-being.

Income and consumption

76. The first recommendation of the SSFC Report is "to refer to income [per capita] and consumption rather than GDP." Even though we strongly believe that documenting the intensity of production provides important information to the general public and policy makers alike, we agree that for capturing changes in material well-being, one should also report changes in income and consumption.

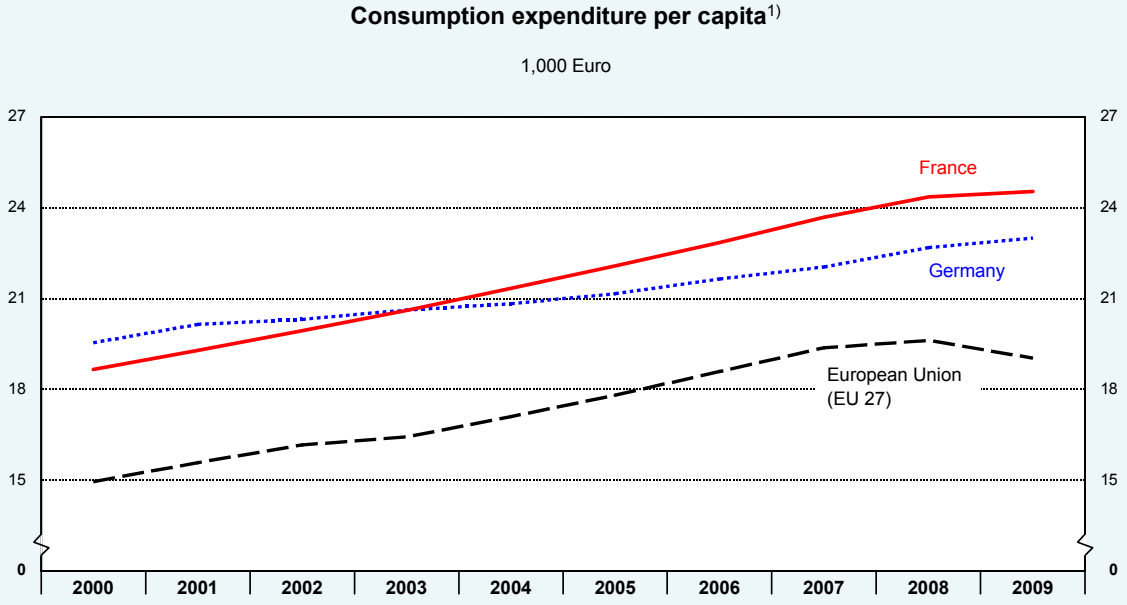
If we want to measure a nation's income, i.e. the income of all its domestic economic agents, the best indicator is **net national income** (NNI) per capita. For countries like France and Germany, NNI moves in close accordance with GDP. The pattern is quite different for countries with large cross-border factor income flows or large inward and outward investment

flows such as Ireland. While for many industrial countries NNI per capita is closely correlated with GDP, it could arguably be considered to be the best indicator of material well-being of national economic agents and should consequently be included in our dashboard.

77. Alternatively, one might also focus on **household disposable income** per capita, **household consumption** per capita or **total consumption** per capita, as the SSFC Report proposes in its second recommendation. As, however, we ultimately want to retain only a small number of indicators in our dashboard, we should adopt only one of them. Household consumption is the indicator more closely linked to the utility function that should summarize individuals' aspirations. The difference between household disposable income and household consumption is household savings. The saving rate is obviously a key parameter of the economy, but it is more relevant to growth sustainability and will consequently be dealt with in chapter IV.

Households' **housing** purchases are regarded as investments and hence are not captured directly in consumption. However, in the national accounts, a housing expense is computed for home-owners and added to rents paid by tenants, making it possible to treat consumption as inclusive of all material needs. Another potential disadvantage of choosing consumption as a measure of well-being concerns the treatment of publicly provided **in-kind services**. Efforts have been made to filter out in-kind services provided by government, such as health or education. However, many government expenses, such as security and justice, aim to ensure households' well-being. We therefore suggest that the consumption indicator should consist of the sum of **household and government consumption** (Chart 8). It should be expressed per capita for the same reasons set out above regarding GDP.

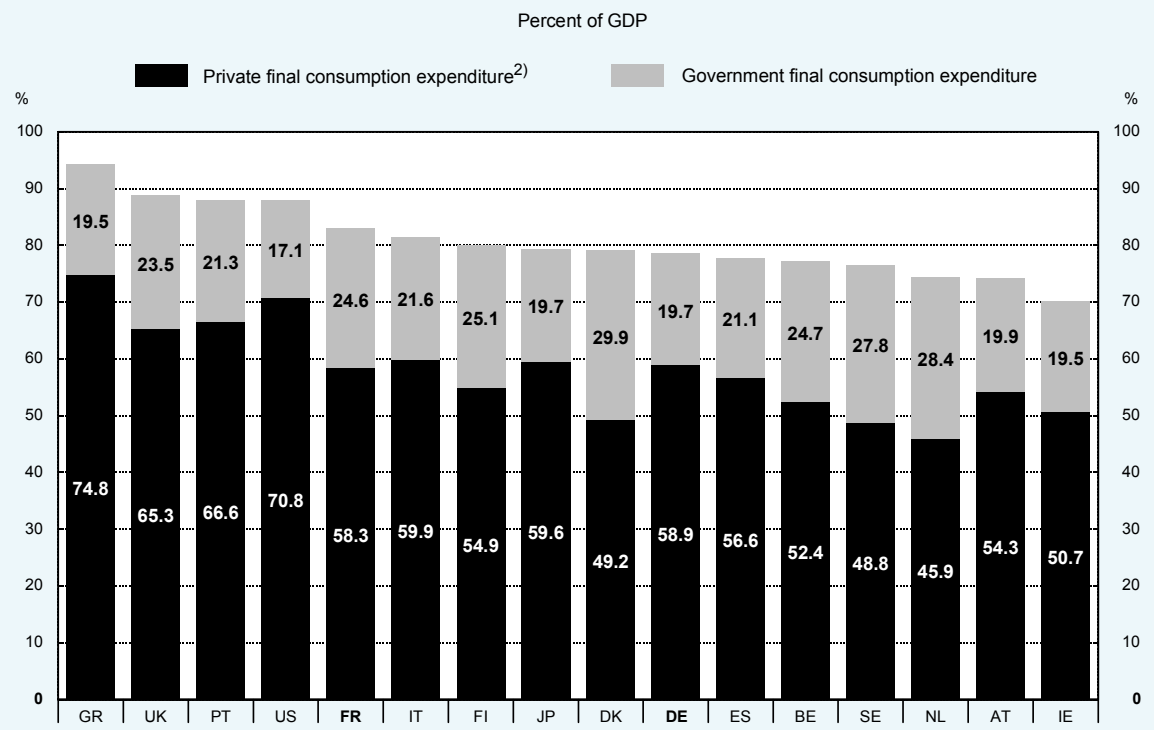
Chart 8



78. The **international comparison** of breakdowns of final consumption between households and governments reveals wide differences that are at least partly due to different national choices in social policy. The sum of the two aggregates ranges between 70 % and 90 % of GDP Among those countries that consume more (and save less), some experience more public consumption through goods and services provided by the government (northern European countries or France), while others feature stronger household expenditures (United States, Japan). This demonstrates that the saving rate cannot be considered a direct consequence of the size of government (Chart 9). We propose to select **final consumption expenditure per capita** for our dashboard, as it captures government consumption largely dedicated to households, although we are aware that some inefficiencies are attributed to governments and can hamper international comparisons of well-being.

Chart 9

Share of private and government consumption expenditure in selected countries (2009)[1)]



1) GR-Greece, UK-United Kingdom, PT-Portugal, US-United States, FR-France, IT-Italy, FI-Finland, JP-Japan, DK-Denmark, DE-Germany, ES-Spain, BE-Belgium, SE-Sweden, NL-Netherlands, AT-Austria, IE-Ireland.– 2) Households and non-profit institutions serving households (NPISH).

Data

Source: EU

Income distribution

79. A meaningful assessment of progress in material well-being can hardly rest on reporting average or median income, but rather needs to take **distributional issues** into account. Correspondingly, recommendation number 4 of the SSFC Report emphasizes the importance of considering the distributional characteristics of income. This request is even more pressing in a world in which inequalities tend to be increasingly pronounced. Specifically, some studies have shown that most income growth in the U.S. in recent years has been captured by a handful of the **highest-income households**. Both in France and Germany, the situation appears to be more complex, as the lowest income-holders have also benefited from powerful redistribu-

tion mechanisms via taxes and subsidies. As a result, the **middle class** has been **squeezed** by these simultaneous alterations happening at both ends of the income distribution.

80. In recent years, both in France and Germany, individual-level data have become available which comprise detailed information on taxes and social transfers, thus facilitating a detailed analysis of the income distribution. In particular, it has therefore become possible to compare income distribution before and after social and tax transfers. A recent study in France even includes the consideration of in-kind services for every income quintile. This has yielded a very informative picture. In the transformation from the distribution of primary incomes to **adjusted** (for in-kind services) **disposable incomes**, about 10 % of primary income is redistributed from the two richer quintiles to the two poorest. A large share of this redistribution operates through the provision of in-kind services (Table 4).

Table 4

Household income distribution in France: 2003[1)]

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
			%			Bln Euro
Primary income	5	12	17	24	42	1,140.2
Disposable income	8	13	17	22	40	993.4
In-kind transfers	25	21	19	18	18	229.5
Thereunder:						
Health	21	22	21	18	19	97.8
Education	28	20	19	18	15	75.1
Housing	70	23	5	1	1	10.2
Adjusted disposable income	11	15	17	21	36	1,222.9 [a)]

1) Individual households residing in metropolitan France, excluding FISIM.- a) Disposable income and in-kind transfers.
Source: INSEE

Data

It would be very attractive to conduct these analyses in an internationally comparative fashion. It is quite **difficult**, however, to find a **harmonized** statistical indicator across many countries. At EU level, there is a common survey based on a household panel that can provide comparative information (EU-SILC). Because of its size – large but not large enough – the results can be calculated at best for population quintiles. For the time being, the results emerging from this exercise should be treated very cautiously, especially when engaging in cross-country comparisons.

81. Measuring the income distribution comprehensively is one important challenge, finding a parsimonious representation which captures its essence is another. In fact, there are many ways to condense the income distribution into a summary figure.

- The most general is the **Gini index**, as it applies to the overall distribution. For the most part, however, the calculation of the index is not easily understood, and it requires detailed and extensive information on the entire distribution, including the highest incomes. The method consists in comparing the actual income distribution with a hypothetical distribution in which everybody has the same income. In this hypothetical distribution, the Gini index is zero. In the other extreme, if all income is held by only one person, the Gini index equals unity.

– The **at-risk-of-poverty** rate is often used to qualify disparities in income distribution. But it only provides information on individuals with a very low income. This indicator is defined as the share of persons with an equivalized disposable income below the risk-of-poverty threshold, which is frequently set at 60 % of national median equivalized disposable income (after social transfers). The indicator is specifically dedicated to quantifying poverty risks but tells us nothing about the upper part of the income distribution.

– The indicator easiest to calculate and understand is the income ratio between the x % with the lowest and the x % with the highest income. When x equals 20 %, the ratio is called **S80/S20**. This ratio of total income received by the top quintile to that received by the bottom quintile of the income distribution is regularly computed by Eurostat (Chart 10). Here

Chart 10

Disparities in income distribution

Income quintile share ratio (S80/S20)

1) PT-Portugal, GR-Greece, UK-United Kingdom, ES-Spain, IT-Italy, EU 27-European Union, EA 16-Euro Area, DE-Germany, IE-Ireland, FR-France, BE-Belgium, NL-Netherland, FI-Finland, AT-Austria, NO-Norway, DK-Denmark and SE-Sweden.

Source: EU

Data

as well, "income" must be understood as equalized disposable income. We suggest including the S80/S20 ratio in our dashboard as it is the easiest to obtain and to communicate to the general public.

82. In all these analyses the definition of **reference income** is important. Many studies on inequality simply look at wages, as they are the easiest statistics to obtain. For our purpose, however, it is necessary to take a much broader view by including all other sources of income besides labour income, preferably on the basis of a comprehensive data base, such as tax statistics. To move from the household to the individual level, incomes should be calculated per capita using an appropriate **equivalence scale** (e.g.1 for the first adult, 0.5 for other adults and children above 14, and 0.3 for other children) so as to incorporate any economies of scale associated with household formation.

In the EU, statistical offices jointly implement a specific survey that already yields information on distribution (EU-SILC). Owing to differences in sample size between countries in EU-SILC, only the quintile distribution can be determined with sufficient reliability. Using this data one can, for instance, calculate the S80/S20 ratio. Correspondingly, we suggest including Eurostat's **income quintile share ratio** in our dashboard. In any event, there is an obvious need to invest in **larger surveys** as well as in the use of **tax sources** in order to improve these data and to be able to calculate a much more precise distribution of income per unit at EU level.

Wealth and time allocation

83. The SSFC Report has also pointed out that income is not the only component of material well-being. Wealth should be considered as well, as it captures the **capacity** of economies to generate material well-being in the longer run. In that sense, it ought to be viewed as a major factor affecting the sustainability of well-being, an issue which will be discussed in more detail in chapter IV. Measuring wealth is difficult, however, especially at an individual level. Consequently, the role of wealth and wealth distribution in material well-being has so far received insufficient attention. While we appreciate the impact that wealth might exert on consumption and investment behaviour, due to this **dearth of information**, we do not propose selecting any wealth indicator for our dashboard.

A first obstacle to the inclusion of wealth in the analysis of current material well-being is the difficulty of measuring **asset prices**. In particular, equity prices are subject to wide fluctuations, and depend heavily on whether they are measured before or after a bubble burst. Furthermore, as the recent crisis has shown, housing prices can be equally volatile. In short, the measurement and valuation of wealth at macro level are challenging tasks – to say the least. Second, individuals tend to think of their wealth as **confidential information**, which greatly hampers the collection of data at the individual level.

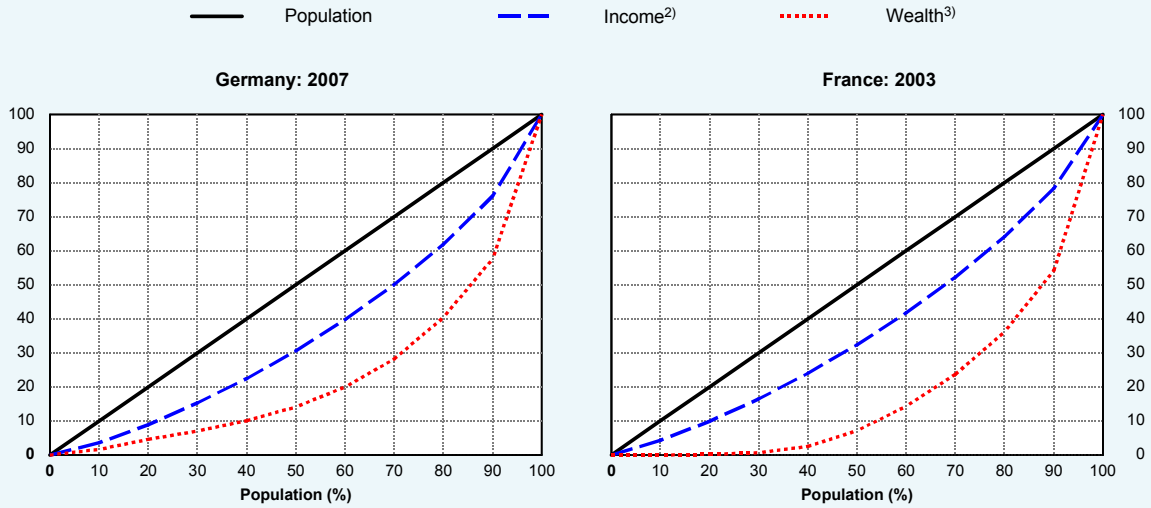
84. Moreover, as wealth tends to be **heavily concentrated**, surveys must be well-stratified and not too small in order to capture the high variance in different types of capital. France, for instance conducts a specific survey (Enquête Patrimoine) every five years. The latest

one (2009) focused on the wealthiest strata but the results are not fully available. Similarly, an analysis of wealth distribution is conducted every five years on the basis of the German Socio-Economic Panel (GSOEP). A further promising approach in this field is the wealth census conducted by the System of European Central Banks.

According to the available empirical results, wealth seems far more concentrated than income (Chart 11). In France, the top tenth ("decile") of capital owners possessed approximately 46 % of total wealth in 2003; by contrast, the top decile on the income scale earned only approximately 22 % of total income. Due to the various conceptual and empirical difficulties involved in analyzing wealth, the data for Germany are not directly comparable to those figures (Chart 12).

Chart 11

Cumulated distribution of income and wealth in Germany and France[1)]



1) Data are not comparable across countries.– 2) After redistribution.– 3) Total household assets. German data include pension entitlements.

How to read this chart: In the case of France, for instance, 80 % of the population receives 64 % of total income after redistribution and owns 36 % of total household assets.

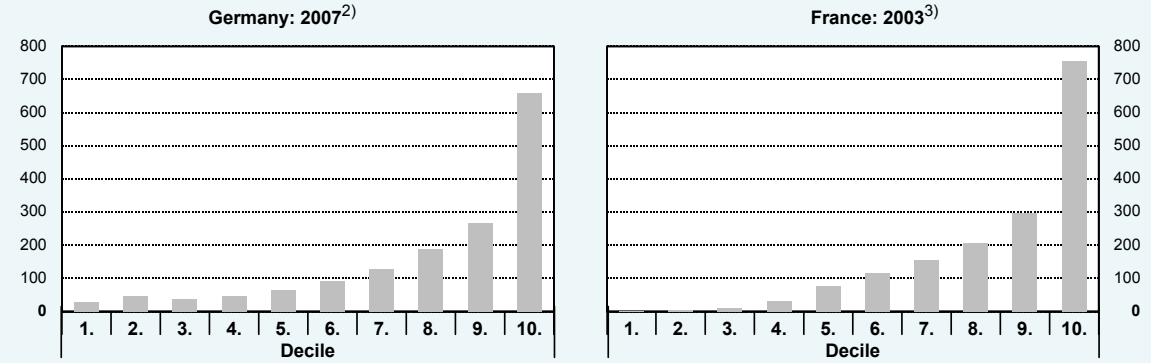
Data

Sources: INSEE, SOEP

Chart 12

Average household wealth per decile in Germany and France[1)]

1,000 Euro



1) Data are not comparable across countries.– 2) Including pension entitlements.– 3) For details see Household Wealth Survey (2003).

Data

Sources: INSEE, SOEP

In particular, they include imputed values for the net present value of expected pension payments, which disproportionately influences the estimated wealth at the lower end of the wealth distribution. Nevertheless, the German wealth data convey the same general message of a highly skewed wealth distribution. For the year 2007, for instance, the corresponding figures are approximately 42 % for total wealth and approximately 24 % for income, respectively. A more **dramatic** reflection of wealth **concentration** is the ratio of the wealthiest decile of the population to the least wealthy decile. This ratio is merely 1 to 5 for income after redistribution for France and 1 to 6 for Germany, respectively, but its value tends to increase by up to two orders of magnitude when looking at wealth data (Les revenues et le patrimoine des ménages 2010; German Socio-Economic Panel 2007).

85. Similar problems as those experienced for wealth, namely that information is irregular and international comparisons are tricky, occur in connection with the issue of **time use**. That is why, despite its undoubted importance for a completely comprehensive assessment of material well-being – and despite the fact that the fifth recommendation of the SSFC Report suggests more frequent time-use surveys – we are not pursuing this issue much further at this juncture. Unfortunately, in our assessment, time-use surveys are still **too infrequent** to produce a suitable indicator for our dashboard. By contrast, the construction of satellite accounts on wealth and time use is a worthwhile undertaking, and research on these topics should be encouraged emphatically.

Intermediate conclusions

86. To sum up this section, we have introduced indicators on income, consumption, and wealth, emphasizing their distributions on both a household as well a per-capita basis. This discussion has led us to the conclusion that the following three indicators should be selected for the dashboard:

- net national income per capita,
- final consumption expenditure per capita including government consumption, and
- a harmonized distribution measure of net income per consumption unit, S80/S20.

5. Concluding remarks

87. The present chapter has reviewed the five first recommendations of the SSFC Report. Its first recommendation is the request to assess the current state of **material well-being** on the basis of income per capita and consumption rather than GDP, which nevertheless remains a valuable indicator of economic performance. Second, the SSFC Report recommends emphasizing the household perspective when material well-being is at issue, while the third recommendation alerts researchers to consider wealth as an important facet of material well-being. A fourth recommendation of the SSFC Report emphasizes the importance of distributional characteristics of income, consumption and wealth, and, finally, a fifth recommendation suggests broadening the perspective to include non-market activities.

Our discussion has been guided by the insight that, although there is always scope for augmenting material well-being further, for **wealthy societies** such as those of France and Germany, it is already an achievement to maintain the existing high level of productive activities. Thus, monitoring economic performance remains an important task, and implementing refinements of GDP which serve this task even better is an important objective for economic and statistical research. Nevertheless, the SSFC Report reminds us of the need to be aware of the limitations of GDP as a measure of well-being, a theme which has been discussed by economists for many decades. Therefore, our report has explored promising avenues for proceeding from the measurement of economic performance to an assessment of material well-being.

88. Most decision-makers would certainly like economists to provide them with "the" ultimate indicator of material well-being. We fully agree with the overarching conclusion emerging from the SSFC Report that this idea is totally unrealistic. In order to proceed from this fundamental insight towards the practical implementation of more realistic alternatives to previous statistical reporting practices, we propose **six indicators** which seek to strike an appropriate balance between comprehensiveness regarding economic performance and the current state of material well-being, on the one hand, and parsimony, on the other. These indicators are:

- GDP per capita,
- GDP per hours worked as a measure of economic productivity,
- employment rate for the 15-64 age group,
- net national income per capita,
- final consumption expenditure per capita, including government consumption,
- an internationally harmonized distribution measure of net income per consumption unit (income quintile share ratio S80/S20).

89. We have also proposed **concrete steps** that need to be taken rapidly – notably the harmonization of panel data on household income – to facilitate consistent measurement of changes in income distribution, such as the EU-SILC (Survey on Income and Living Conditions) panel. In particular, the sample size should be expanded if we want to gain more comprehensive knowledge not only of differences in income distribution but also of other factors linked to well-being. Regular studies comparing time use across countries should also be considered. Furthermore, we have outlined the need for further statistical advances in fields such as in-kind services and intangible activities – and, more generally, in the statistical coverage of various economic sectors.

Reforming the system of indicators of economic performance and current material well-being is important. But to effectively develop a new compass for policy-making, the crucial step will be to anchor communication on progress to a system of indicators that takes better account of non-material sources of well-being and the sustainability of current modes of behavior and levels of well-being. These issues are addressed in the following chapters.

Appendix: Breaking down macroeconomic figures to take account of disparities among households

Two types of information are available on income and consumption in France: the household account and household surveys. The method used at INSEE consists in building a bridge between the two approaches in order to break down household accounts by household category.

Specifically, the "household accounts" are broken down using data from the national accounts for 2003 and from five INSEE surveys on income and consumption. Statistics on Household Income and Living Conditions (EU-SILC), Taxable Income, Household Budget, Housing, and Health. From these individual data, the aggregate macroeconomic totals for income or consumption are allocated to the different categories of households.

Each component of disposable income and consumption expenditure (wages and salaries, benefits, rent, etc.) is broken down by household category in the following stages.

- Identification of the survey that would provide the closest definition to the one used by the national accounts for the component examined (e.g. for health expenditure, the Health Survey rather than the Household Budget Survey).
- Calculation of the average amounts for each household category (e.g. average salary for each standard-of-living quintile).
- Calculation of the associated financial totals by multiplying the average amounts by the number of persons in each category.
- Re-adjusting the overall totals obtained to the totals in the national accounts, the coverage of which is confined to ordinary households in mainland France.

Each component of disposable income and consumption expenditure determined from the national-accounts totals is thus broken down by household category. This facilitates the deduction of total disposable income by aggregating all income components for a given category. The same procedure is conducted for consumption expenditure. Finally, savings and saving rates are deduced on the basis of these breakdowns.

To compare disposable income and consumption expenditure by category, the total figures are first divided by the number of households in the category, and then by the average number of consumption units in the category.

Transfers between household residents are included. By definition, when total household accounts are compiled, financial transfers (such as maintenance payments and financial assistance) and exchanges of goods and services (such as cars, clothing, and electrical goods) between resident households are therefore neutral for accounting purposes and are not evaluated separately. However, these transactions are not distributed uniformly among households. Private transfers are made mainly to young people. Therefore they have to be estimated and taken into account in the household breakdown. Total amounts were taken from the House-

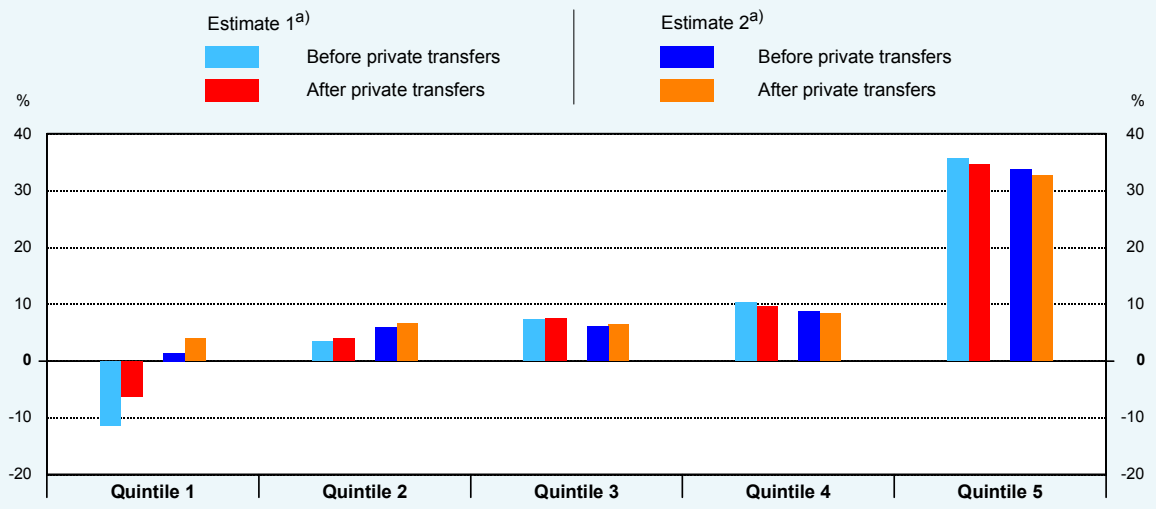
hold Budget Survey, as was the breakdown of average amounts received and paid per category.

These studies will form the basis for measuring the change in purchasing power in each household category and hence the change in inequality between households in the national accounts. In France, the figures for 2003 showed that the disposable income of the wealthiest 20 % (quintile) of households was five times that of the least wealthy 20 %. More than half of the lowest quintile's income consisted of social benefits, and one-third of its spending went on expenses described as "precommitted" – such as rent and housing charges, telephone services, and insurance – which are hard to negotiate in the short term.

The primary expenditure item for all households, whatever their living standard, is housing. It represents on average almost one quarter of the consumption budget, a proportion that increases with age. Working and retired households in the same category have fairly similar living standards, but the elderly consume less, while broadly maintaining the consumption habits of their earlier years.

Chart 13

Saving ratio per income level in France: 2003[1)]



1) Individual households residing in metropolitan France, excluding FISIM.– a) For explanation of definitions see Family Budget Survey (2006).

Data

Source: INSEE

The saving ratio – the fraction of disposable income that is not consumed – increases with the living standard and with age. It is particularly high for the self-employed, whose saving may also be directed toward maintaining and improving the tools of their trade. Overall, the most prosperous households save over one third of their income. Conversely, the most modest households generally cannot save at all. Their saving rate is even negative, estimated at between -11 % and 1 % in 2003. The main beneficiaries of private transfers among households are young people, single-parent families, and the least well-off: their saving rate is higher after monetary transfers have been taken into account.

References

Atkinson, T. (2005) *Atkinson Review: Measurement of Government Output and Productivity for the National Accounts*, Final Report, London: Palgrave MacMillan.

Braibant, M. (2006) *Compiling Input-Output Tables in France since 1950,* mimeo.

Cutler, D., A. Deaton and A. Lleras-Muney (2006) *The Determinants Of Mortality*, Journal of Economic Perspectives, 20 (3), 97 - 120.

Fleurbaey, M. and G. Gaulier (2007) *International Comparisons of Living Standards by Equivalent Incomes*, CEPII Working Paper, 2007 - 03.

Fournier, J.-M. and D. Marionnet (2009) *Measuring Banking Activity in France*, Banque de France, Quarterly Selection of Articles, No. 16, Winter 2009 - 2010.

INSEE (2005) *Économie française, 2005 - 2006,* http://www.insee.fr/fr/themes/document.asp?reg_id=0&ref_id=ECOFRA05d.

--- (2009) *L'économie française – Comptes et dossiers*, ed. 2009, http://www.insee.fr/fr/themes/document.asp?ref_id=ECOFRA09d.

--- (2010) *Les revenus et le patrimoine des ménages*, Insee Références series, 2010 ed., http://www.insee.fr/fr/publications-et-services/sommaire.asp?codesage=REVPMEN10.

Nordhaus, W. D. and J. Tobin (1973) *Is Growth Obsolete?*, NBER Chapters, in: The Measurement of Economic and Social Performance, National Bureau of Economic Research, Inc., 509 - 564.

CHAPTER III

Quality of Life

Quality of Life

90. There is more to life than material well-being. Mankind would indeed be poor if all which we were striving for was material in nature. Therefore, this expertise acknowledges the **diversity** of **human existence** and, when describing the current state of affairs and the most recent developments, deliberately goes beyond the documentation and discussion of indicators of material well-being. When it comes to implementing this decision, however, we have to make a fundamental choice: should we amalgamate our indicator of material well-being with supplementary, non-material information into an **overall measure** of "happiness"? We clearly consider this to be the **wrong avenue**, due to problems of interpersonal and intergenerational comparability, possibly overwhelming errors of measurement and misperception, and, finally and most importantly, the manipulability of such a measure.

Instead, we insist on taking our fundamental idea of offering a **dashboard** for orientation seriously. An informed society should be confronted with the diversity of life by a **set of indicators** which strike an **appropriate balance** between providing sufficiently rich information and preventing its recipients from becoming overwhelmed by its complexity. In our assessment, there is simply no sensible alternative to supplementing the indicators of material well-being developed in the previous chapter by a parsimonious set of indicators of the current state of key non-material aspects of life. The principal idea is that any **weighting** of such aspects has to remain a completely **idiosyncratic affair**. In that sense, all we are offering is the truth. Nothing more.

1. Conceptual questions: the blue pill or the red pill?

91. Two fundamental approaches can be taken to capturing the abstract notion of quality of life: a "top-down" approach – i.e. moving from an overall measure of subjective well-being to its constituent elements – or a "bottom-up" approach – i.e. starting from individual facets of human existence and then gravitating towards a comprehensive assessment of well-being. After carefully considering the respective **conceptual arguments**, we clearly advocate adopting a bottom-up perspective, since the empirical implementation of the top-down approach rests on very strict **identification assumptions that** in our assessment are unpalatable.

Top-down approaches: seductive, but unconvincing

92. At first sight, it might seem that we could meet our key objective of reporting the current state of affairs fairly easily by taking a **top-down perspective**. This perspective would acknowledge that all reasoning about human welfare concerns the **subjective well-being (SWB)** of individuals, and not simply certain objective facets of well-being like income or consumption. In that sense, material indicators are poor proxies for the genuine concept. If it were indeed possible to **directly measure** "contentment" or "satisfaction with life" or "happiness" in an empirically plausible way, then, according to this reasoning, there would be a hope of using this information in the construction of an **aggregate** quality-of-life **indicator**. This metric would then be ready to replace objective measures like gross national product (GNP) altogether.

As such an indicator would summarize and aggregate the SWB of each individual member of society, its **principal elements** would be the individuals' subjective overall assessments of their **levels of well-being** ("top"). For purposes of reporting the state of affairs and its most recent developments, one would not need additional information regarding any constituent aspects of the SWB experienced by individuals. An example of an attempt to construct such an indicator is "satisfaction with life", surveyed in various waves of the World Values Survey. The correlation of this measure with GNP per capita, though clearly positive, is far from being perfect (Chapter I, Chart 1). Thus, proponents of this measure will feel vindicated in doubting the capability of GNP to serve as an encompassing indicator of welfare.

93. As a consequence of its very stringent **identification assumption**, namely that "The genuine SWB can be ascertained by direct observation", the most important challenge for the top-down perspective would then be mastering the measurement of SWB on an individual level. Researchers would have to tackle **problems** of misjudgement and strategic misrepresentation by respondents. Routines would have to be developed to collect information on a regular and timely basis and at a reasonable cost. Finally, it would be important to ensure comparability across different societies and points in time. All these obstacles would only be **technical** in nature and, thus, would seem comparatively easy to overcome. And yet, the whole approach rests on a central premise which we do **not** find particularly **convincing**, namely the direct observability of genuine SWB.

94. If one were convinced, however, the **political relevance** commanded by such a measure would be twofold. First, government **performance** could be assessed directly by comparing overall SWB at different points in time. Second, knowing how those aspects of life that policies can influence shape overall SWB could become important for **policy design**. As a consequence, in addition to assessing the level of SWB the top-down approach would then require the factors that influence SWB to be identified and their effects quantified in a supplementary analytical step.

Potential **determinants** of life satisfaction are plentiful. For instance, Layard states the "Big Seven Factors affecting happiness", among them family relationships, the financial situation, health, and personal freedom (Layard, 2005). Similarly, Frey and Stutzer distinguish personality, socio-demographic, economic, contextual, and institutional factors as important determinants of SWB (Frey and Stutzer, 2001). Much of the research in this literature has been concerned with the **correlation** of overall indices of SWB with these factors. These correlations have been derived on the basis of a number of empirical approaches, most importantly using survey data.

In any case, the **fundamental assumption** underlying this line of reasoning is that SWB can be measured directly and explained empirically. The effect of becoming unemployed, for instance, can then be directly compared with the effect of getting divorced (Table 5). For those determinants of well-being which could be manipulated by governments, say "quality of government", one could use similar analyses to deduce **policy prescriptions** and assess the effects of policy measures.

Table 5

Effects on happiness[1)]

	Fall in happiness (points)
Financial situation	
Family income down by a third	2
Family relationships	
Divorced (rather than married)	5
Separated (rather than married)	8
Widowed (rather than married)	4
Never married (rather than married)	4.5
Cohabiting (rather than married)	2
Work	
Unemployed (rather than employed)	6
Job insecure (rather than secure)	3
Unemployment rate up 10 percentage points	3
Community and friends	
"In general, people can be trusted"	
Percentage of citizens saying yes down by 50 percentage points	1.5
Health	
Subjective health down 1 point (on a 5-point scale)[2)]	6
Personal freedom	
Quality of government	
Belarus 1995 rather than Hungary 1995[3)]	5
Personal values	
"God is important in my life"	
You say no to this rather than yes	3.5

1) Source: Layard (2005) and Helliwell (2003), based on World Values Survey. In three waves of this survey, 87 806 people from 46 countries reported life satisfaction on a scale from 1 to 10 (mean 6.8; standard deviation 2.4). For this table, the effect of each single item is estimated, keeping other features of life unaltered. Estimation is least squares with „fixed effects" for country groups, waves, age, education, and societal variables. The numbers are multiplied by ten, so that the life satisfaction scale is between 10 and 100.– 2) Reading: Keeping everything else constant, a fall of subjective health (measured on a 5-point scale) by one point, happiness falls by six out of 100 points.– 3) Reading: Keeping everything else constant, a move from Hungary to Belarus in 1995 reduces happiness by five points.

Data

95. And yet, despite the euphoria with which these advantages are often described, on closer inspection the underlying idea of direct measurability of individual SWB is **unacceptable**. Most importantly, the top-down perspective has to be grounded in the assumption that the subjective measures of well-being are in fact a **truthful reflection** of a **genuine state** of well-being which the individual is aware of but which – due to the complexity of affairs – cannot be fully grasped by the researcher. It is only under this condition that the "hard" indicators which to date have taken centre stage can convincingly be replaced by a more encompassing measure of the kind described here.

We believe, however, that there is ample reason to follow the presumption of most empirical work in economics, namely that facts speak louder than words, and that nothing reveals genuine preferences more transparently than actual choices made. Statements about preferences will always be an imperfect or even misleading surrogate for such **acts of revelation**, as the

translation from the genuine state of well-being to the statement recorded by the researcher is influenced by many intervening factors, such as by strategic response behaviour.

96. Moreover, the top-down perspective conflicts with existing evidence on **discrepancies** between facts and perception. Many people would fail to realize that a major improvement in their quality of life has taken place during the previous decades, even though value added and the related consumption opportunities, as well as many other objective factors, have indeed increased or improved. If such misperceptions prevail, it seems to make little sense to construct measures of well-being or even to formulate any policy prescription based on such subjective statements. Thus, if one accepts discrepancies between facts and perception as an **important facet** of human existence, the top-down approach quickly loses its appeal. Simply administering the top-down approach and establishing a measure for SWB on the basis of survey data would then be inadequate. In our view, there are already enough attempts to conceal the actual state of affairs in a bold pursuit of self-interest. Thus, resources should rather be allocated to collecting objective information, improving its transparency and intensifying its propagation.

Bottom-up approaches: reasonable, but challenging

97. Consequently, it seems better to approach the problem of non-material influences on the quality of life from an angle that does not rely on the strict identification assumption embodied in the top-down perspective. For example, the **capabilities approach** (Sen, 1999) focuses on the functionings and the freedom of a human being, acknowledging the difficulty of amalgamating these different elements of well-being into a single indicator. We advocate following such a more demanding avenue and taking a distinct **bottom-up perspective**. The starting point of this bottom-up perspective is – again – the recognition that a range of different factors make life worth living and that only a subset of these factors can be valued in monetary terms.

This approach, however, starts from the individual non-material aspects of human existence and then gravitates "up" towards overall well-being, instead of from a (rather imperfect) measure of well-being "down" towards its constituent elements. The strategy behind this second, our favoured, approach is that if it were possible to (i.) organize the abundance of relevant factors systematically along a limited number of **dimensions** and to (ii.) condense factors at least within each dimension to an operational **indicator**, a more accurate picture of social welfare would emerge, even though we would **not** attempt to use these elements to **construct** a single **encompassing measure** of social welfare. The components of this comprehensive picture would have to be weighted by the **recipients** of the information, instead of by the researchers as its **producers**. After all, this is the very idea behind reporting a dashboard of relevant information which is the recurrent motive of our report.

98. Implementing this approach throws up three challenges. First, it is necessary to **systematically organize** the plethora of different factors in "**dimensions**". This organization has to balance the desire of reflecting the complexity of the phenomenon with the need for parsimony. Correspondingly, we define "dimensions" as **groupings of indicators** which address

similar aspects of human existence such that the resulting dimensions could not justifiably be condensed any further without considerable loss of information. The question of deciding where to draw the line in applied work inevitably involves some **ambiguity**. We think that the considerations of the Stiglitz-Sen-Fitoussi Commission with respect to this demarcation problem are an appropriate starting point. For instance, for implementing the capabilities approach Nussbaum (Nussbaum, 2000) also discusses ten dimensions, among them bodily health, emotions, and affiliation, a list of entries which – as our discussion in the next section of the chapter will demonstrate – is quite similar to that drawn up in the SSFC report.

99. Second, **individual indicators** that are suitable for describing each of these dimensions comprehensively have to be selected from a very large number of potential measures. To accomplish this task, two conditions have to be fulfilled. On the one hand, the selected individual indicators need to illuminate as completely as possible the **full range** of experiences related to a dimension. For example, in the area of health, not only mortality rates for children but for all relevant subgroups of the population have to be included. This implies that the correlation between indicators must not be too high. On the other hand, the number of indicators should remain **small enough** to be manageable. Specifically, the additional knowledge gained by adding an indicator should lie above a pre-specified threshold.

Whenever possible, it makes sense to rely on "hard" individual indicators, which are (i.) collected on a regular and timely basis (ii.) at reasonable cost and (iii.) are comparable over time and across societies. Even so, **"hard" indicators** will certainly not suffice to capture all facets discussed in their entirety. They need to be complemented by other indicators. The first toehold for measurement is **social indicators**, the collection of which was broadened significantly in the 1970s. The European System of Social Indicators contains several hundred measures in various life domains. However, from the perspective of the capabilities approach they mostly measure achieved functionings or states, but not opportunities. In order to obtain data for capabilities in the latter sense, **additional information** is needed.

100. The third challenge is then the proper condensation of the individual indicators into a respective **overall indicator** within each of the dimensions. To this end, it will be necessary to define **aggregation weights** which seek to capture the value individuals assign to the various aspects of each dimension. These must either be derived from empirical studies (via contrasts, "up") or set by the researcher based on a priori reasoning or statistical analysis. In any case, one must always invoke a range of more or less stringent **identification assumptions** to condense the available information into an overall indicator. Consequently, any hope of finding a unique solution for this aggregation problem would be futile.

In principle, economists have a lot to say about preference orderings. A prominent empirical approach rooted in **welfare economics** is the concept of willingness to pay. A society may be found indifferent, for instance, between an average annual income of 30,000 euro with an average life expectancy of 75 years, on the one hand, and the combination of 55,000 euro and 65 years, on the other. In that case the welfare gain of an increase in life expectancy might be attributed a monetary value. Yet their willingness to pay depends strongly on the existing in-

come of agents and hence tends to be biased towards the better-off. Alternatively, the theory of **fair allocation** tries to remedy the shortcomings of the willingness-to-pay concept by finding reference sets of individual situations which allow for a comparison of welfare among individuals. Of course, finding the right reference set might be just as problematic as measuring individual preferences. Consequently, in practical work the assignation of monetary equivalents is a daunting task.

101. It is the key premise of our dashboard approach that the overall indicators which will be constructed for each dimension should **not** be **aggregated further** into an **overall indicator** of the quality of life. After all, the defining feature of each dimension is that further aggregation would only be possible at the cost of large reductions in information content. For instance, one could perhaps justify, after careful consideration, the concept of aggregating the mortality of different age groups, the use of preventive measures and the typical waiting time for a medical treatment into an overall "health" indicator. But we consider the idea of balancing "health" concerns with "social participation" to be conceptually unjustified, as the non-cardinal nature of the individual indicators and the heterogeneity of individual preferences preclude the provision of a meaningful aggregation scheme by the social researcher.

102. The condensation of individual-level information into comprehensive indicators necessarily raises three further cross-cutting issues. First, one has to ask how **inequality** can be taken into account, since averaging over members of a society will always imply a loss of information about the distribution of the phenomenon. If inequality is considerable, a focus on the population mean might conceal a serious societal problem. With regard to material well-being, this concern was addressed in the previous chapter. As this contribution is a first attempt to seriously augment our regular reporting on society's well-being with non-material facets, it seems nevertheless advisable to concentrate on **first moments** of their distributions. In our view, it will only be advisable to break further ground and analyze **higher moments** of these distributions once this enhanced strategy has passed the test of time.

Second, characteristics such as income, education, and health may be highly **correlated** across the various dimensions of quality of life and, thus, the interactive nature of deprivations will be ignored by analyzing only one dimension at a time. Instead, one would have to analyze **joint distributions**. Yet, for the same reasons that convince us to pursue first moments exclusively, we would be hesitant to move beyond **marginal distributions** for the time being. Finally, inequalities might **persist** over time, thereby precluding the existence of equal opportunities. Snapshot analyses cannot identify such problems. While this is a highly relevant issue for social research, the nature of year-to-year reporting unequivocally suggests retaining a strict focus on current information.

2. Empirical implementation: a tough task

103. Within each dimension of quality of life many different **facets**, captured by **individual indicators** as their operational equivalent, contribute to the state of affairs. In line with our bottom-up strategy, we need to condense the ample information provided by the complex universe of all available indicators into a more palatable **overall indicator**, thereby deliberately

incurring a loss of information. Recent econometric research has outlined interesting avenues for achieving this objective empirically with an eye on reflecting individual heterogeneity (e.g. Ferrer-i-Carbonell and Frijters, 2004; Frijters et al., 2004). Yet, for the time being we advocate relying instead on a priori reasoning or on statistical approaches to complexity reduction.

Addressing heterogeneous preferences

104. In economic reasoning, **individual preferences** take centre stage. Unfortunately, as they are not observable directly, it is far from straightforward to implement this focus empirically. This also impedes the analysis of individual indicators of quality of life, since individuals' appreciation of life's variegated facets are far from obvious. In principle, researchers can take one of two basic approaches to gaining a satisfactory description of the valuations attached to different circumstances: they can rely on answers in **survey data** or on the **implicit revelation** of preferences by actions. Traditional empirical economics certainly favours relying on approaches of revealed preference, following the insight that it is comparatively realistic to seek to capture **preference orderings** from the contrasts of choices made between different bundles of goods. However, the hope that one might be able to capture the **utility level** associated with a pre-defined bundle of goods without strict identification assumptions is slim. If one wants to determine that utility level or measure the appreciation of aspects of life not traded on markets, one is forced to largely rely on the analysis of surveys.

105. Unfortunately, survey data are often plagued by significant deficits. (A fortiori, these observations hold true for implementing direct measures of overall SWB in accordance with the top-down approach.) Misunderstandings or misperceptions on the part of the respondents or simply negligence may lead to severe **measurement errors**, and answers might even be distorted by respondents engaging in **strategic responses**. Moreover, short-term positive and negative effects influence subjective well-being alongside overall life satisfaction, and these **short-term emotions** might contaminate survey information on life satisfaction. Finally, answers might be distorted by **partial adaptation** to a new situation, for example to a suddenly occurring disability or to the fulfilment of an individual desire ("hedonic treadmill").

Even if these problems can be neglected, major problems of comparability over time and across societies emerge. It is likely that answers differ merely because of the **social context**, even if factual circumstances are held constant. Whether certain events are perceived as having a drastic and lasting impact on life satisfaction tends to depend on a society's view of the degree of their inevitability. A fortiori, reference points seem to be important for understanding reported subjective well-being (Helliwell and Barrington-Leigh, 2010).

106. In the recent **econometric literature**, serious attempts are undertaken to rid survey data of such subjective aspects, for example by making use of insights from **panel data econometrics**. Typically, these analyses employ multiple observations of the same observation unit and invoke identification assumptions corresponding to this data situation. Assuming that the unobserved cultural imprint of response behaviour, a form of "unobserved heterogeneity", is constant over time, the marginal effects of variations in the explanatory variables can be suc-

cessfully identified. While this would enable observers to analyze contrasts, levels of well-being would still be difficult to measure. This limitation also holds when researchers use **alternative concepts** of measurement, for example brain imaging. Again, it is far from obvious how this should help to identify the level of well-being, and yet this is our ultimate objective.

Purely statistical approaches

107. In lieu of convincing empirical work providing us with a sufficient range of preference-based aggregation weights, for a given dimension one might rely on a priori reasoning and select one particular **headline indicator** out of the reservoir of candidate indicators as its representative indicator. This route to overcoming the intricate problems of measurement involved in the attempted discovery of preference orderings is both **robust** and **transparent**. Since it is easy to understand and interpret, and can be applied even in situations where only very few indicators exist at all, it is widely used in practice. Nevertheless, it will always be very difficult to justify a particular selection based on a priori reasoning. In effect, even the most carefully chosen headline indicator will always be suspected of being **highly subjective** or even arbitrary.

Still, their robustness makes the use of headline indicators a very strong candidate approach, and we deliberately apply it extensively in our own application. We mainly employ headline indicators in situations where adequate indicators are **rare** and in which they are highly **correlated**. In many dimensions characterizing quality of life, only a few suitable indicators are measured at all. Moreover, some indicators are only constructed at wide or even irregular **intervals**. And, finally, many of the individual indicators might not be internationally comparable.

108. A closely related alternative to selecting a single headline indicator is the choice of a **composite indicator** readily available from a statistical office or a research institution. Typically, they are constructed as linear combinations of individual indicators. While this approach might often be subjective as well, depending on the weighting scheme devised by its supplier, it **detaches** the **construction** of the overall indicator from its **application** to analyzing the state of well-being. Most importantly, relying on the experience and knowledge of those organizations that have set up such existing composite indicators may frequently be regarded as a step towards objectivity.

109. A further alternative is the application of a **statistical approach** to complexity reduction. The statistics literature has suggested a multitude of procedures for aggregating complex information (OECD, 2008). They all share the objective of retaining as much as possible of the information contained in a large set of variables, subject to the restriction that it is represented by a reduced set of variables. In line with our bottom-up strategy, we are aiming to construct just one overall indicator within each dimension which represents the state of affairs there as accurately as possible. Following such an algorithm has the advantage of being **more objective** than selecting a headline indicator could be. The obvious disadvantage is the **mechanical nature** of any such procedure, since algorithms do not pay any attention to the con-

tent of those variables whose information they are compressing. Therefore it seems advisable to apply statistical approaches with considerable caution.

110. Our statistical approach of choice to complexity reduction is **Principal Components Analysis** (PCA), a simple, non-parametric method. PCA seeks to reduce the dimensionality of a data set consisting of a large number of interrelated variables while retaining as much as possible of the variation present in the original data set (Jolliffe, 2002). In principle, one could always construct several principal components, but our aim is to construct just one (the "first") principal component, which is then used as the overall indicator for the dimension under scrutiny. This component is a **weighted average** of the underlying individual indicators and captures as much of their variance as possible. Using the first principal component for comparisons of levels **across countries** would be **problematic**, because weights are different and the proportion of the variance that is explained by the first principal component differs as well. Therefore we refrain from doing so. Nonetheless, we will use PCA extensively to compare the development **across time**, albeit separately for France and Germany.

Since PCA proceeds from the definition of a set of individual indicators, some element of subjectivity is inevitably involved in the collection of this initial set. Consequently, there will simply **never** be a **completely objective** approach, and the results of a statistical approach to complexity reduction will always depend on data availability and on the competence of the empirical researchers applying it. In particular, serious problems might arise if important individual indicators are omitted from the initial set of variables because they are not collected at all or only at irregular intervals. On the other hand, including superfluous information in the initial set is equally problematic. Therefore, we use the Kaiser-Meyer-Olkin (KMO) measure as a **formal test** to check whether the variables selected have enough in common overall to warrant a PCA analysis (Kaiser, 1970, 1974).This measure takes values between 0 and 1, with small values indicating insufficient commonality. More concretely, we base our decision on how to proceed on a threshold value of 0.5 for the KMO statistic.

111. In descriptive statistical analyses, PCA is the method of choice whenever the number of individual indicators is substantial. In our application to dimensions of quality of life, the collection of a **large** and **comprehensive** set of individual indicators is a critical first step. PCA loses its appeal, though, when this preliminary work is not successful or when a priori reasoning or the desire to provide easy-to-communicate results strongly suggest the choice of a headline indicator. In addition, in practical work PCA does not always lead to robust and plausible results. Therefore, we decided to pursue **both approaches** simultaneously, with PCA serving as a testing tool for the relevance of the chosen headline indicators. After all, headline indicators are also constructed as a weighted average of individual indicators. But while PCA derives the corresponding weights according to a pre-specified **algorithm**, the researcher selects one individual indicator to receive the weight unity and all others to receive a weight of zero. In the ideal case, the headline indicator and the first principal component will have so much in common that it becomes irrelevant which approach should be publicly communicated. In other cases, the researcher has to make a conscious choice.

3. Practical implementation: France and Germany

112. As all theory is grey, and green the golden tree of life, we have accordingly applied our proposed bottom-up strategy to the actual reality of **two countries**, France and Germany, for **three exemplary years**, with the year 2000 as the anchoring year. In a first step, we have selected a range of eight **dimensions** of quality of life, only one of which is concerned with material living standards. In this endeavour, we have benefited tremendously from the insightful work of the Stiglitz-Sen-Fitoussi Commission. In a second and third step, we have collected extensive information on **individual indicators** derived from a large set of sources, and constructed **overall indicators** for each dimension. This section contains a concise overview of our results, while the subsequent section provides a comprehensive account of our detailed analyses.

Choosing dimensions

113. The **Stiglitz-Sen-Fitoussi Commission** has taken up the task of differentiating quality-of-life dimensions in due consideration of both subjective well-being and the capabilities approach, i.e. in our terminology the top-down and the bottom-up perspective, respectively. According to our assessment, it seems advisable to interpret their suggestions from the **bottom-up perspective**. The Commission has identified **eight dimensions** of interest, including material living standards, that are mostly objective and lean towards the capabilities and fair allocation approach and against subjective well-being. The dimensions chosen are displayed in the first column of Table 6. While the first dimension, material well-being, is captured – for better or for worse – in the system of national accounts and was addressed in the previous chapter, the other seven dimensions together form the central topic of this chapter. Moreover, the penultimate dimension on the list, environmental conditions, leads on to the next chapter on sustainability.

114. With only one overall indicator representing each dimension of quality of life, it is all the more essential that the chosen dimensions **completely** cover all relevant aspects of quality of life. Furthermore, **measurability** and **political relevance** should be taken into account. Under these supplementary criteria, the categorizations of Nussbaum as well as Frey and Stutzer do not appear very appealing (Nussbaum, 2000; Frey and Stutzer, 2001). Given our preference for objective measures of the quality of life, the same applies to the proposal of Layard (Layard, 2005). From a practical perspective, there is necessarily a high degree of overlap between the dimensions proposed by the SSFC and other official bodies such as the OECD: thus some use “financial situation” or “economic factors” instead of material well-being, while health dimensions are alternatively labelled “bodily health” or “physical and mental health” (Giovannini et al., 2009).

If one accepted that environmental conditions are satisfactorily dealt with under the heading of sustainability, and that personal and economic insecurity is a cross-cutting issue of previous dimensions, one could stick with the six dimensions chosen by the OECD. As, however, “insecurity” as well as “environment” have a direct influence on the quality of life we would rather not go down this route, as in our opinion this would imply too big a loss of information.

Overall, we believe that the dimensions chosen by the Stiglitz-Sen-Fitoussi Commission provide the **ideal balance** between comprehensiveness and focus.

115. The second column of Table 6 provides some examples of the **facets** that **contribute** to each of the eight dimensions of quality of life. This collection has to serve as a **starting point** of the search for individual indicators, the second step of our empirical strategy. A detailed discussion of each dimension can be found in chapter 3 of the SSFC report. For the sake of brevity, we do not intend to replicate it here in its entirety.

Table 6

Quality of life – dimensions and its facets

Dimensions	Examples of contributing facets
Material well-being	Income, consumption, change in wealth, income distribution
Health	Life expectancy, diseases, disabilities, infant mortality, physical and mental illnesses, health distribution
Education	Basic reading and writing skills, knowledge of calculus, problem solution competence, information and communication technology, pupils and students performance, life-long learning, education distribution
Personal activities	Working, commuting, various kinds of recreational activities, distribution of personal activities
Political voice and governance	Citizens' voice, legislative guarantees, rule of law; possibility to participate in the political process, voter turnout, membership rates of parties, unions, non-governmental organisations, participation in protests, degree of democracy, independence of media, corruption, distribution of political voice
Social connections and relationships	Family relationships, friends, intensity of friendships, social contacts, distribution of social connections
Environmental conditions	Availability of clean air, water and soil, reachability of pleasant environment next to one's home, climate, distribution of environmental conditions
Personal and economic insecurity	Risk of illnesses, injuries, damages, theft, robbery, murder, death, unemployment, social exclusion, becoming poor, distribution of personal and economic insecurity

Data

This collection is a rich reservoir of contributing facets which arguably cover the spectrum of human experience quite comprehensively. **Material well-being**, one of the topics dealt with in the previous chapter, is the outcome of income, wealth and consumption. **Health** is to capture both data on life expectancy and on the prevalence of diseases. **Education** serves both as a direct source of well-being and an indirect one: skills and knowledge help to intensify positive experiences. What people do each day clearly influences their quality of life. Apart from sleeping, all kinds of working and of leisure activities contribute to the dimension of **personal activities**.

The dimension **political voice** and **governance** serves to capture the contributions of a well-functioning democratic society to quality of life. A related topic, addressing individuals and families more directly, is the well-functioning of **social connections** and **relationships**. Next to personal and societal circumstances, the quality of the four elements that surround us is

subsumed under **environmental conditions**. Uncertainty about the future and corresponding fears often constrain quality of life. The intensity of this effect depends on the degree of **personal** and **economic insecurity** which is considered as the last dimension in this chapter.

Quality of life in France and Germany

116. On the basis of these carefully selected dimensions, we have systematically applied the second (finding appropriate individual indicators) and the third (constructing an overall indicator for each dimension) steps of our strategy to the cases of France and Germany. It is the purpose of this exercise to pave the way for future regular reporting. Consequently, in the process of **choosing individual indicators** for each of the dimensions, diverse aspects such as regular availability, timeliness, coverage, comparability and reliability have been taken into account. Overall, we have been **quite successful** in finding at least one individual or even composite indicator which captures important facets of the respective dimension and satisfies these supplementary conditions. However, this search has **not always** led to a **comprehensive** set of indicators that span the complete spectrum of facets contributing to this dimension. A more thorough account of our implementation of this second step in the concrete application to France and Germany is provided in the following section.

In the next step of **determining** an appropriate **overall indicator** for each dimension, we have followed two routes whenever possible. In any case, our discussion has identified one individual (or composite) indicator which might arguably represent the state of affairs in this dimension as its **headline indicator**. Moreover, to cross-check our choice, wherever the initial data set of individual indicators permitted, we have **constructed** overall indicators by employing **PCA** separately for France and Germany. Yet this procedure can only lead to convincing results if the various facets of the respective dimension are represented comprehensively by a rich set of variables which have been collected consistently over a long period of time. PCA is certainly not the right method for tracking progress across countries.

117. The **seven overall indicators** chosen to represent the **non-material dimensions** of quality of life in our concrete application are presented in Table 7. As the dimension **material well-being** was described in detail in the last chapter, it is largely disregarded in this chapter. If one were to include a single indicator, though, this would probably be net national income per capita. Some of the non-material indicators listed in Table 7 still have to be implemented, such as the composite indicator for education, while others are not (yet) published annually or with a sufficiently short time lag. But these indicators would be our **favourites** for **future regular reporting** of the state of well-being, and their use could easily be ascertained by appropriate policy decisions.

118. It goes without saying that our analysis is open to **improvement** by future research. We would greatly welcome and encourage pertinent, constructive comments by the scientific community as well as practical men and women. In particular, the results of our PCA used for cross-checking are far from perfect, due to relatively short time series and the problem of omitted variables. Moreover, in an annual presentation of the dashboard the PCA would have to be repeated each year, yielding **new weights** for the underlying variables. Any **new data**

points should be added, thus yielding longer time spans, so as to improve reliability. Furthermore, any **new variables** collected should be added as well to alleviate possible problems of omitted variables.

Table 7

Proposed quality-of-life indicator

Dimensions	Proposed indicator
Material well-being	See chapter 2
Health	Potential years of life lost (PYLL, OECD), to be replaced by healthy life years (HLY, Eurostat)
Education	Students aged between 15 and 24 as a percentage of the population of the same age group (Eurostat), possibly to be replaced by Programme for the International Assessment of Adult Competencies (PIAAC, OECD)
Personal activities	Employees working on shift work (Labour Force Survey)
Political voice and governance	World Bank Institute Worldwide Governance Indicator "Voice and Accountability"
Social connections and relationships	Frequency of spending time with people at sport, culture, communal organization, World Values Survey 1999/2000
Environmental conditions	Urban population exposure to air pollution by particulate matter (Eurostat)
Personal and economic insecurity	Not-at-risk-of-poverty rate (SOEP, Eurostat), possibly to be replaced by Personal Security Index (to be developed in accord with that of the Canadian Council on Social Development)

Data

119. Ultimately, the **complete empirical results** regarding economic performance, material well-being, quality of life and sustainability should be reported together, as exemplified in chapter one of this study. In years to come variants of this table (the "**dashboard**") could routinely list the absolute values of all chosen indicators and their respective changes from the previous period. In this chapter, we go beyond this mundane documentation, however, and visualize the results regarding the seven **non-material quality-of-life** dimensions in the form of **radar charts** for France and Germany. In the radar charts, an increase in the indicator implies an improvement in the respective dimension (Chart 14).

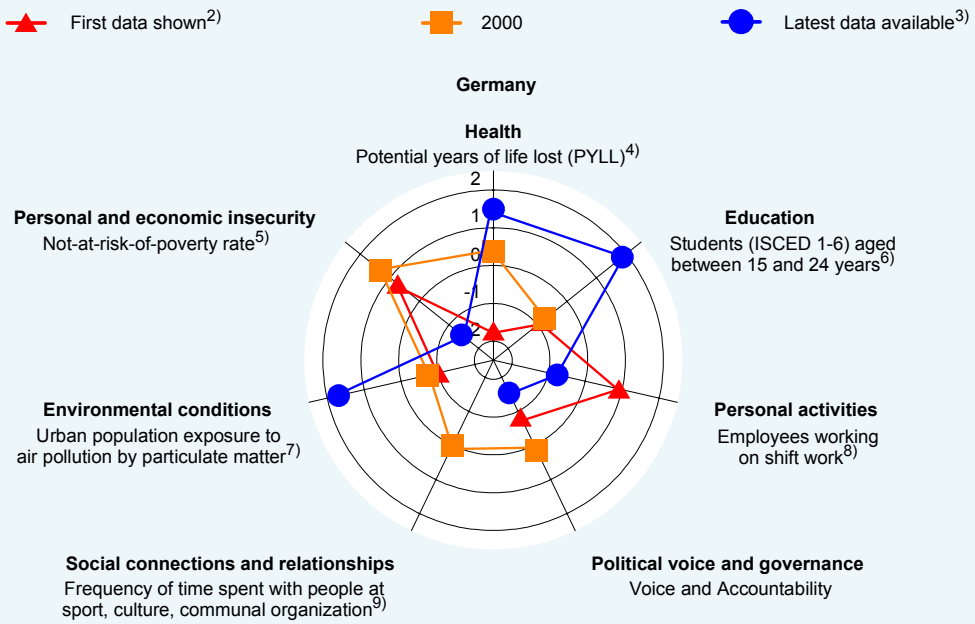
For any given set of years, these country-specific radar charts separately report the **evolution** of the two societies along each of the dimensions. For reasons of consistency one would usually compare a fixed set of years. In our application, however, the time dimension is chosen to depict the evolution for the longest available **period** for each of the overall indicators. The year 2000 serves as an anchor for all indicators, however. In addition, results are presented for the earliest and most recent years possible.

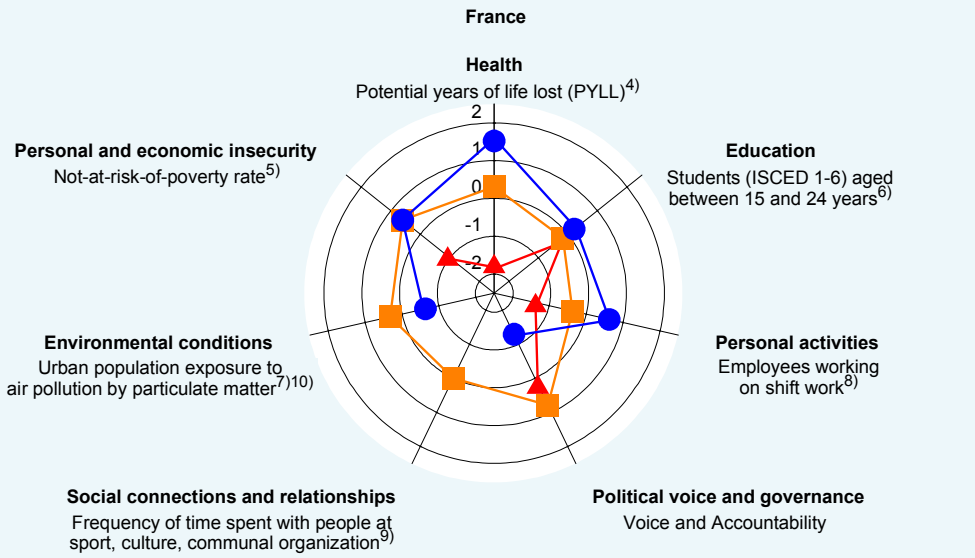
For purposes of presentation, the overall indicators of each dimension are typically **normalized**, and therefore their absolute values do not have any obvious direct interpretation. Here the indicator values are calculated by subtracting the mean and dividing by the standard deviation. This normalization has been conducted for each country individually. Thus, **country comparisons** – which at any rate are not justified for the case of PCA – are **not possible**.

Moreover, as tempting as this visual representation might make it seem, the **surface** of the radar chart could **never** be a **valid** measure of the overall quality of life, since this would imply the completely unwarranted equal weighting of its dimensions.

Chart 14

Non-material quality-of-life indicators[1)]





1) Own calculations; values are not comparable across countries. Average = 0; value higher than 0 implies better conditions and vice versa.– 2) Health: 1991, Personal activities: 1992, Political voice and governance: 1996, Education: Germany: 1992, France: 1993, Enviromental conditions: Germany: 1999, France: 2001, Personal and economic insecurity: Germany: 1992, France: 1995 .– 3) Health: 2006, Education and Personal activities: 2009, Political voice and governance and Enviromental conditions: 2008; Personal and economic insecurity: Germany: 2009, France: 2008.– 4) PYLL is a summary measure of premature mortality which provides an explicit way of weighting deaths occuring at younger ages, which are, a priori, preventable. In relation to 100,000 population, calculated by the OECD Secretariat based on age-specific death statistics provided by the World Health Organization.– 5) One minus share of persons with an equivalised disposable income below the risk-of-poverty threshold, which is set at 60 % of the national median equivalised disposable income after social transfers.– 6) In relation to the population in the same age group.– 7) The indicator shows the population weighted annual mean concentration of particulate matter at urban background stations in agglomerations.– 8) As a percentage of total employees.– 9) Only data available: 1999.– 10) For 2000: 2001 data.

Sources for calculations: EU, OECD, SOEP, The World Bank, World Values Survey

Data

120. The first non-material dimension of quality of life, **health,** is represented by **potential years of life lost**. This indicator collects information on premature deaths, sums up the differences between age at death and 70 years for each of them, and presents the result per 100,000 people. The data base is very reliable and time series span a long period. It shows a steady improvement for both France and Germany over the past decades. Nonetheless, this indicator does not capture data on the prevalence of diseases. In our assessment, the optimal headline indicator to capture both mortality and morbidity would be healthy life years (HLY), as collected by Eurostat. As soon as consistent and reliable data are available for a sufficient intertemporal comparison, we propose to switch to this indicator.

121. To obtain an indicator for **education** we currently rely on students aged between 15 and 24 years as a percentage of the population of the same age group, as collected by Eurostat. While data for Germany show a steady increase, the share is decreasing in France. One should keep in mind that this series does not capture a measure of education output (skills), but rather of schooling output (graduation), which is not necessarily coincident. We therefore strongly recommend testing adult competencies regularly and frequently. The OECD initiative Programme for the International Assessment of Adult Competencies (PIAAC) can be seen as an ideal source for a regular indicator.

122. No measure yet exists that compounds both work and leisure as the two main branches of **personal activities**, and the association between respective time series is rather moderate. Starting from initiatives on quality of employment and decent work, we regard data on "working hours and balancing work and non-working life" as a valid gateway for this dimension. To select a single headline indicator, we propose to focus attention on the percentage of employees working shifts. The resulting numbers suggest a decline in quality of life due to shift-working in Germany, while in France the situation has improved during the last few decades.

123. As yet there is no existing measure of **political voice** and **governance** through regular surveys. As long as this lack of data persists, we propose as an imperfect substitute the World Bank Institute's Worldwide Governance Indicator on "Voice and Accountability" that relies mostly on expert opinions as our indicator for this dimension. In worldwide comparisons of this indicator France and Germany are ranked among the top countries in all observation periods.

124. The perhaps least accessible dimension is **social connections** and **relationships**. The only promising approach to this dimension seems to be the analysis of survey questions, and one of these should be posed regularly. We suggest this should be the question "Frequency of spending time with people at sport, culture, communal organization", as included in the World Values Survey 1999/2000. As comparable annual survey results do not exist, intertemporal comparisons of the results are so far not possible.

125. General **environmental conditions** indicators of quality of life are rare, with the notable exemption of air quality. Given our focus on output measures and the lack of existing composite indicators, we select the "urban population exposure to air pollution by particulate

matter" of diameter less than 10 micrometers (PM10) as the headline indicator. PM10 can be carried deep into the lungs where they can cause inflammation and a worsening of the condition of people with heart and lung diseases. According to our empirical results, both countries display a continuous improvement in environmental conditions.

126. Aggregating data for the dimension of **personal** and **economic insecurity** is a daunting task, as this dimension has many facets. We propose relying on a headline indicator, the **not-at-risk-of-poverty rate**. Eurostat defines the at-risk-of-poverty rate as "the share of persons with an equivalized disposable income below the risk-of-poverty threshold, which is set at 60 % of the national median equivalized disposable income (after social transfers)". We calculate the not-at-risk-of-poverty rate as one minus the at-risk-of-poverty rate because an increase in this rate implies an improvement of the situation. As we are aware that this indicator does not capture too well all the facets of this dimension, we propose to evaluate in future research the possibilities of replacing this indicator either by a PCA indicator (subject to the condition that time series with sufficient properties exist) or by an index obtained similar to the Personal Security Index of the Canadian Council on Social Development. According to the not-at-risk-of-poverty rates, France has experienced a rather stable development in this quality-of-life dimension, whereas Germany improved slightly until around 2000 and worsened since then.

127. In sum, while health and environmental conditions have unequivocally improved over time and in both countries, the documented developments with respect to education, to personal activities, to political voice and governance, and to personal and economic insecurity provide a more variegated portrait of societal progress throughout the two recent decades.

4. Elements of our dashboard: a detailed discussion

128. In our exemplary dashboard, for the sake of brevity we have presented empirical results for each quality-of-life dimension for at most three years. In this section we present a thorough discussion of our indicator choices for readers interested in the detailed research underlying these summary statements. Most importantly, we consider possible choices for headline indicators within each quality-of-life dimension under scrutiny. And whenever possible, we also provide detailed results of the PCA and compare these results with these possible headline indicators so as to highlight their strengths and weaknesses.

Health

129. Health is probably the most fundamental quality-of-life dimension, as a lack of health usually has a negative impact on all of the other relevant dimensions. It is therefore not surprising that national and international organizations provide a multitude of health indicators, though coverage is often diverse. A large set of **individual indicators** is concerned with **mortality**, such as median life expectancy or life expectancy at birth, while many other indicators capture aspects of **morbidity**. Most importantly, morbidity indicators comprise information on the prevalence of different diseases, measures of self-reported health and anthropometric measures on height and weight. Specific indicators such as infant mortality, the life expec-

tancy of different age groups, death rates due to particular chronic diseases, the prevalence of obesity or smoking, the incidence of serious accidents at work regularly provide indispensable information for experts and policymakers. However, for the purpose at hand, they are too narrowly focused on single conditions or groups of the population.

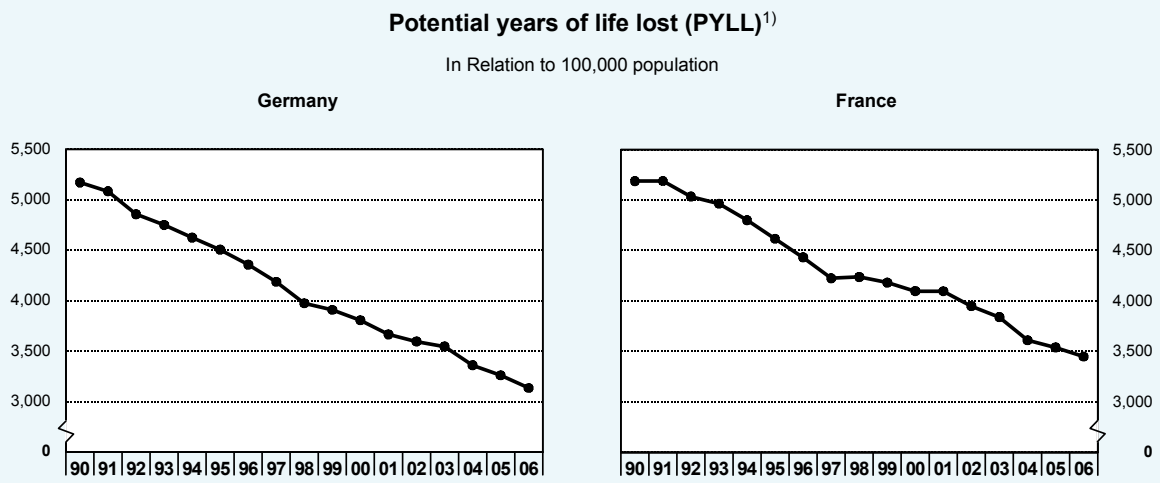
130. Among the more comprehensive indicators, measures of life expectancy are typically the first demographic numbers considered. Yet, they disregard any negative effects which disease or disability might exert on the quality of life. To circumvent these problems, **combined indicators** of health conditions merge mortality and morbidity data into a single indicator. Several candidate indicators of health status have been suggested in recent years. For example, many relevant aspects can be captured by the concept of disability-free life expectancy. The **healthy life-years** (HLY) indicator builds upon the remaining years a person of a certain age can expect to live without disability, and thus combines information on infant mortality, the prevalence of disabilities and the life expectancy of adults into a single indicator: the healthy life years that a newborn can expect to live given the prevailing conditions.

Alternatives to the HLY indicator, such as **disability-adjusted life years** (DALY) or **health-adjusted life expectancy** (HALE), usually face the challenge of requiring weights which express the valuation of different health states. For example, to calculate HALE, a measure of the number of years spent in full health, each year is multiplied by a weight that is high for minor diseases and low for diseases that impair physical and mental functioning more severely. In the case of death, the value of a year is taken to be zero, and in the case of full health, it is counted with weight one. Similarly, in the case of DALY, premature death, prolonged illness and disabilities are subtracted from potential life years, where one year of life with a specific illness or disability is assigned a certain percentage of a year in full health. Given the complexity of the weighting scheme and the difficulties of weighting across different cultures, we take the view that a simple **binary choice** – as is necessary for HLY – is preferable for representing the quality-of-life dimension health.

131. **Eurostat** has been collecting information on HLY since 1995, following the European Union Sustainable Development Strategy (EU SDS) which defines a "healthy condition" by the absence of limitations in functioning. Thus, episodes of illness without limitations in functioning are counted as disability-free periods of life. While the **mortality** data used to construct this indicator are typically of high quality and comparability, data on the respective proportions of the population with and without **disabilities** were until recently only imperfectly comparable over time and across countries, as survey instruments and methodologies were not synchronized. In addition, cultural differences might influence the answers regarding questions on certain disabilities. From 1995 to 2001 these data were obtained from the European Community Household Panel. In the years 2002 and 2003 existing data were extrapolated. Then followed a period of transition to the new EU Statistics on Income and Living Conditions (EU-SILC). From 2004/05 onwards these data have been based on standardized surveys within EU-SILC, enhancing their **comparability** over time and across countries. If current procedures are kept in place, HLY thus promises to provide a reliable and intertemporally as well as internationally comparable headline indicator.

132. Yet, for the time being, we need an **interim headline indicator** for the dimension health, at least as long as only a few data points exist for the desired HLY indicator (especially for Germany), making it impossible to gauge its reliability over a lengthy time span. In our application, we suggest using **potential years of life lost** (PYLL), a weighted mortality indicator which has been reported by the OECD for a couple of decades. This indicator collects information on premature deaths. For each person who dies below the age of 70 the difference between age at death and 70 years is retained, and the total number of these potential years of life lost within one calendar year is then related to 100,000 people. This indicator shows a nearly steady improvement both in France and in Germany (Chart 15).

Chart 15



1) PYLL is a summary measure of premature mortality which provides an explicit way of weighting deaths occuring at younger ages, which are, a priori, preventable. Calculated by the OECD Secretariat based on age-specific death statistics provided by the World Health Organization.

Data

Source: OECD

133. In addition to this interim headline indicator PYLL we present the result of a **PCA** for the dimension health, using "OECD health data" for the period 1996-2006 in the case of Germany and 1993-2004 in the case of France. Since we intend to only use data whose ordering with respect to health is clear, we do not use data on health expenditures or on the numbers of employees or graduates in the health sector, although these are common individual health indicators. We also drop data on the prevalence of different diseases because of the intricate problem of accounting for competing risks.

As our initial set of variables we employ information on the subtopics **prevention** (vaccination rates against measles and DTP (diphtheria, tetanus and pertussis)), **mortality** (life expectancy at birth and at 65 years old, infant mortality and potential years of life lost) and **mental problems** (suicides). A higher share of preventive protection and a higher life expectancy both at birth and at 65 years of age arguably document an increase in the health situation of the population. Therefore the sign of these variables' weights should be positive. A rise in potential years of life lost (our interim indicator) and suicides is expected to express a negative development of the health status of the population and, thus, the sign of this indicator's weight should be negative. According to our descriptive analyses, each variable that is expected to indicate an improvement in the health situation of the population has increased over

time. Similarly, each variable that tends to indicate a worsening of the health situation has decreased for both countries. Therefore the weighted average obtained by PCA is strongly expected to increase (Table 8).

Table 8

Health – Variables for the principal component analysis[1)]

	Germany		France	
	1996	2006	1993	2004
Vaccination rates against measles, % of children immunised	*86.6*	*94.5*	*78.0*	*87.1*
Vaccination rates against DTP, % of children immunised	*94.1*	*97.4*	*95.0*	*98.0*
Life expectancy at birth, females (in years)	80.1	82.4	81.4	83.8
Life expectancy at birth, males (in years)	73.6	77.2	73.3	76.7
Life expectancy at age 65, males (in years)	14.9	17.2	15.9	17.7
Potential years of life lost (PYLL), all causes, females, years[2)]	2,945	2,212	3,079	2,361
Potential years of life lost (PYLL), all causes, males, years[2)]	5,741	4,044	6,861	4,879
Suicides, deaths per 100,000 population	12.4	9.1	18.6	15.0

1) Source: OECD.– 2) PYLL is a summary measure of premature mortality which provides an explicit way of weighting deaths occurring at younger ages, which are, a priori, preventable. The PYLL in relation to 100,000 population are calculated by the OECD Secretariat based on age-specific death statistics provided by the World Health Organization.

Data

Preliminary testing (low KMO values) suggests that one should omit life expectancy for females at 65 years of age and infant mortality from the PCA, since they display a very high correlation with the rest of the variables, rendering their information obsolete. According to our results, the **signs** of each weight used for composition of the first principal component are as expected for both France and Germany (Table 9). As a robustness check we performed PCA with different time sub-samples and alternative variable selections. The results are robust to these changes. For Germany the first component explains 93 % of the **variance** in the data set and for France 88 %, respectively. The overall KMO value is above 0.6 for Germany and above 0.7 for France and therefore high enough to warrant a PCA analysis.

Table 9

Health – Weights of the first principal component[1)]

	Germany	France
Vaccination rates against measles, % of children immunised	*0.417*	*0.398*
Vaccination rates against DTP, % of children immunised	*0.410*	*0.378*
Life expectancy at birth, females (in years)	0.302	0.292
Life expectancy at birth, males (in years)	0.332	0.308
Life expectancy at age 65, males (in years)	0.343	0.314
Potential years of life lost (PYLL), all causes, females, years	– 0.325	– 0.361
Potential years of life lost (PYLL), all causes, males, years	– 0.332	– 0.373
Suicides, deaths per 100,000 population	– 0.351	– 0.388
Kaiser-Meyer-Olkin measure of sampling adequacy	0.613	0.743
Eigenvalue of first principal component	4.910	5.288
Proportion of variance explained by first principal component	0.930	0.880

1) Calculations based on OECD-Data.

Data

134. In line with our bottom-up strategy, the first principal component derived in the PCA serves as the overall indicator for the dimension "health". For both countries this indicator increases between the mid-1990s and mid-2000s, yielding an **improvement** of the health situation throughout the past decade. The same pattern can be seen in one of the underlying time series, our interim indicator PYLL (Chart 15). Consequently, we are including the results of PYLL into our dashboard with considerable confidence.

Education

135. Apart from its immediate contribution to a high quality of life, education exerts indirect effects, since it enables people to intensify the positive experiences of other dimensions. For example, a higher level of education broadens the scope of personal activities that a person can potentially carry out, is usually associated with higher levels of health and reduces economic insecurities by increasing job stability. Therefore, it is important to capture the **skills** and **knowledge** of a society's members with appropriate individual indicators. As Giovannini et al. (2009) forcefully point out, the focus should thereby be on **output** measures instead of **input** measures like education expenditures. Among output indicators, years of schooling or the percentage of people participating in education and training are problematic candidates, as the quality of the respective forms of education is not known and hence international comparability is not ensured.

136. The best output indicators that capture skills and knowledge are probably obtained through **testing** of **achievements** in literacy and numeracy. While these output measures do exist in quite some detail for younger age groups, coverage of the whole population is more limited. But since we are interested in an indicator for education as a source of current quality of life, the education level of all age groups is relevant. Among the available (composite) indicators that capture a broader sample of the population, those based on the **International Adult Literacy Survey** (IALS) and its successors appear to be the most promising starting point.

At the heart of this endeavour lies the understanding that literacy is not a zero-one distinction between those who can read and write and those who cannot, but rather a continuous, multifaceted phenomenon. Specifically, literacy is defined as the ability to use "printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential" (Kirsch, 2001). The IALS asks a representative sample of people between the age of 16 and 65 to read, understand and interpret various texts, covering **prose literacy** (continuous texts like medicine labels, descriptions, manuals), **document literacy** (non-continuous texts as in figures or tables), and **quantitative literacy** (calculations based on information from continuous or non-continuous texts). The results are ranked on a scale from zero to 500, and five proficiency levels are derived. IALS was conducted in 20 countries in the years 1994, 1996, and 1998.

The IALS was replaced by the **Adult Literacy and Life Skills** (ALL) survey conducted in 2003 and 2006 in a subset of these countries. ALL differs from IALS in its third field. Instead of quantitative literacy, ALL features a **numeracy** scale that covers proficiency in estimation and statistics. Furthermore, it includes a fourth field **problem solving**. The OECD picks up

these developments in its **Programme for the International Assessment of Adult Competencies** (PIAAC) survey. It is projected that the first results of this survey will not be available before the end of 2013 and will include the domains literacy, numeracy, problem solving, and information and communication technologies.

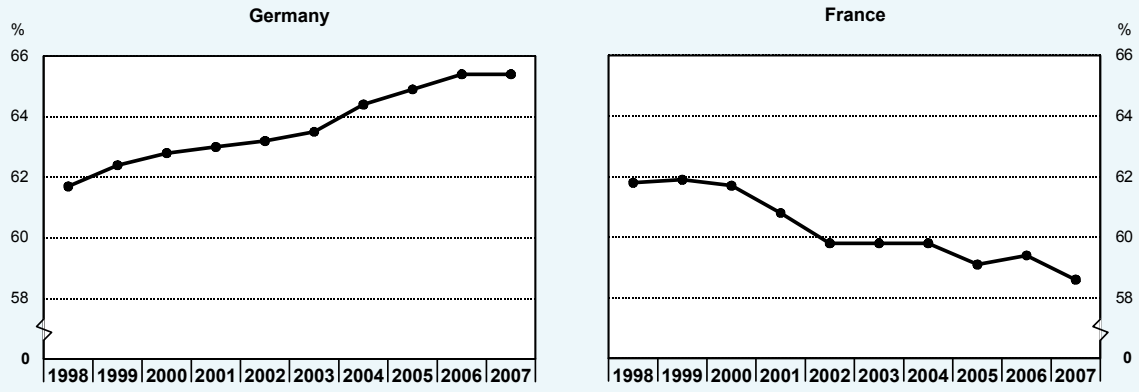
137. Studies based on panel data using skill assessments similar to IALS and its follow-ups document the fact that a lack of skills in the respective domains indeed exerts an adverse effect on many features that are associated with a high quality of life (for example, Bynner and Parsons, 1997). In particular, the positive correlation between low levels of literacy and numeracy and the risk of being unemployed, separated or divorced, physically ill, and less engaged in public activities appears to be robust and rather high. Subject to the condition that the OECD uses an appropriate data collection methodology to ensure reliable information, we propose to present the average scores of the **PIAAC survey** as the **composite indicator** of the education dimension. Moreover, it would be desirable to increase the survey's continuity by carrying it out at least every two years and basing it on a survey design that ensures comparability over time. Reference to associated costs was already made in the first chapter.

138. Until such time as a sufficiently long time series exists, we have to rely on an interim indicator that best serves our purpose. Given our focus on regular reporting and coverage of a broader group of the populace, we propose to use students aged between 15 and 24 years as a percentage of the population of the same age group as an interim indicator. Indicator values are steadily improving in Germany, while values for France show a slight decrease over time (Chart 16).

Chart 16

Students (ISCED 1-6)[a] aged between 15 and 24 years

In relation to the population in the same age group



a) International Standard Classification of Education.

Data

Source: EU

139. In addition to discussing our preferred composite indicator, we conduct **PCA** for the education dimension. Ideally we should use output data that directly measure the increase in skills obtained in the educational system. Yet these data are difficult to collect because the skills of an individual are not directly observable and the available achievement surveys have not been evaluated frequently enough to allow a PCA. Therefore, we have to rely on other

data sets. Specifically, we use Eurostat data for the period 1999-2007 for Germany and 1998-2007 for France. The data cover variables of participation rates, graduation rates and the share of early school leavers (share of individuals aged 18-24 years who have a lower secondary education or less). We use two participation variables: students aged between 15 and 24 years and students aged over 30 years, as a percentage of the respective population of the same age group. And we employ two graduation rate variables: the number of graduates who finished the first or second stage of tertiary education (ISCED 5-6) aged between 20 and 29 years per 1,000 people of the population, and the percentage of the population aged between 25 and 64 years who hold at least a higher secondary school qualification.

Variables capturing the quality of the educational system (output variables) should be used as soon as a reliable data collection procedure is discovered and its data quality is ascertained to be high. In future, output variables from the PIAAC study could be added as further variables to a PCA analysis. The first wave of PIAAC will be available at the end of 2013, but it will take a long time until these variables could be used for PCA because a relatively long time series is needed.

For the variables used in our analysis, an increase of the share of students aged between 15 and 24 years, the number of graduates between 20 and 29 years and the percentage of the population with at least a higher secondary school qualification tends to indicate an increase in the educational level of a society. Thus, the weights of these variables should be positive. For the variable "students aged over 30 years" the direction is unclear, because this group tends to be very heterogeneous. The corresponding weight should be positive when the variable mainly captures mature adults engaging in further education. Conversely, it should be negative if the variable mainly reflects the share of long-term students. Finally, an increase in the share of early school leavers is an indication of a decrease in educational performance and therefore the weight is expected to be negative. According to our descriptive results, except for the share of students aged over 30 years, for Germany the variables indicate an improvement in the educational level. For France the overall tendencies are not that clear-cut because the share of students aged 15 and 24 years decreases (Table 10).

Table 10

Education – Variables of the principal component analysis[1)]

	Germany		France	
	1999	2007	1998	2007
Students (ISCED 1-6) aged between 15 and 24 years[2)]	62.4	65.4	61.8	58.6
Students (ISCED 1-6) aged 30 years and over[2)]	3.3	2.4	1.4	1.8
Total graduates (ISCED 5-6) aged between 20 and 29 years per 1,000 people of the population	31.3	38.6	61.7	77.4
Population aged between 25 and 64 years having completed at least upper secondary education[2)]	79.9	84.4	59.9	68.5
Early school leavers[3)]	14.9	12.5	14.9	12.6

1) Source: EU.– 2) In relation to the population of the same age group.– 3) People aged between 18 and 24 in percent of the population of the same age group who are holding only a lower secondary school qualification and received no further education.

Data

140. As before, we conduct separate PCA for France and Germany and for various sub-samples, achieving sensible and robust results (Table 11). As, for France, the number of graduates aged between 20 and 29 years (ISCED 5-6) per 1,000 people of the population is collected irregularly, the results for France are less reliable than those for Germany. Except for the indicator of the relative share of students aged over 30 years where the direction is unclear, all other signs of the weights turn out to match our expectations. The **first principal component** yields an explanation of the variance of 70 % for Germany and of 93 % for France. According to the KMO value of above 0.65 for Germany and 0.67 for France, the data set warrants a PCA.

Table 11

Education – Weights of the first principal component[1)]

	Germany	France
Students (ISCED 1-6) aged between 15 and 24 years[2)]	0.497	0.552
Students (ISCED 1-6) aged 30 years and over[2)]	– 0.534	0.459
Total graduates (ISCED 5-6) aged between 20 and 29 years per 1,000 people of the population	0.542	0.505
Population aged between 25 and 64 years having completed at least upper secondary education[2)]	0.321	0.391
Early school leavers[3)]	– 0.266	– 0.277
Kaiser-Meyer-Olkin measure of sampling adequacy	0.653	0.673
Eigenvalue of first principal component	3.313	3.834
Proportion of variance explained by first principal component	0.701	0.930

1) Calculations based on EU-data.– 2) In relation to the population of the same age group.– 3) People aged between 18 and 24 in percent of the population of the same age group who are holding only a lower secondary school qualification and received no further education.

Data

141. The first principal component of PCA increases for Germany, as expected, between 1999 and 2007, reflecting the patterns of the individual indicators. Although the inspection of each time series was less clear-cut for France, the overall education indicator obtained by PCA yields an improvement of the educational situation in France between 1998 and 2007. This overall finding is predominantly the same for the proposed interim headline indicator, thus tending to confirm the appropriateness of choosing it.

Personal activities

142. Personal activities are a tremendously heterogeneous dimension. **Time-use surveys** suggest that people pursue very different activities during a day or a week, ranging from commuting and working to spending time in leisure. We can presume that these activities tend to have quite diverse effects on quality of life, but since time allocation at least partially reflects **deliberate choices**, it will be difficult to derive information on the desirability of any specific activity set. A good starting point might therefore be the analysis of indicators characterizing the activity that occupies most of the people for the majority of daylight time – **work**. Work should be seen in this respect not as a means to generate income and thus material well-being, but as an activity whose various facets directly influence quality of life.

The International Labour Office (ILO) (Measurement of decent work), the European Commission (Quality of living and working conditions) and the European Foundation (European working conditions survey) have developed corresponding statistical indicators. The **decent work framework** is a concept that values opportunities for work with the following characteristics. They are (i) productive and deliver (ii) a fair income, while providing (iii) security in the workplace and (iv) social protection for families. Furthermore, they offer (v) better prospects for personal development and social integration, and (vi) freedom for people to express their concerns, they allow them to (vii) organize and participate in the decisions that affect their lives and they provide (viii) equality of opportunity and treatment for all women and men (International Labour Office, 2008). Each of the three indicator sets has the drawback that it suits a particular purpose or policy agenda. Furthermore none of them is broad enough to capture all aspects of quality of employment.

143. Therefore a "Task Force on the Measurement of Quality of Employment" was set up by the United Nations Economic Commission for Europe to define an international conceptual framework for measuring the **qualitative dimension of work** and to propose a set of indicators to measure quality aspects of labour and employment. After several meetings the task forces agreed in October 2009 upon the basic principles of statistically measuring the quality of employment. The framework proposed by the task force is primarily designed to measure quality of employment from the perspective of the worker and not from that of the company. They defined seven substantive elements in their quality of employment framework (UNECE Task Force on the Measurement of Quality of Employment, 2010), namely:

- Safety and ethics of employment,
- Income and benefits from employment,
- Working hours and balancing work and non-working life,
- Security of employment and social protection,
- Social dialogue,
- Skills development and training,
- Workplace relationships and work motivation.

The framework proposed by the task force has been successfully tested twofold. First, nine country reports – Canada, Finland, France, Germany, Israel, Italy, Mexico, Moldova and Ukraine – were prepared by national teams for the task force's final report. Second, a validation study has been conducted by ISTAT used PCA to test the completeness and validity of the proposed indicators (http://www.unece.org/stats/documents/2009.10.labour.htm). The first five elements of the quality-of-employment framework are also included in the decent-work framework, while the last two elements are specific to the former. By contrast, the decent-work framework also addresses employment opportunities (Chernyshev, 2009).

144. The seven substantive elements of the quality-of-employment framework display considerable overlap with other quality-of-life dimensions. To circumvent **redundancy** in our dashboard, we therefore disregard some of these substantive elements in our further analysis.

The first and the fourth substantive elements deal with insecurity, which itself is a quality-of-life dimension. The second substantive element can be considered part of material well-being, and the sixth part of the education dimension, while the seventh element coincides with the dimension "social connections and relationships". Finally, it is difficult to conceive how the fifth substantive element might be relevant for quality of life. So we restrict our attention to the third element "**working hours** and **balancing** work and non-working life". A particular advantage of this element is that it pertains not only to the sphere of work but also to leisure and other non-working activities.

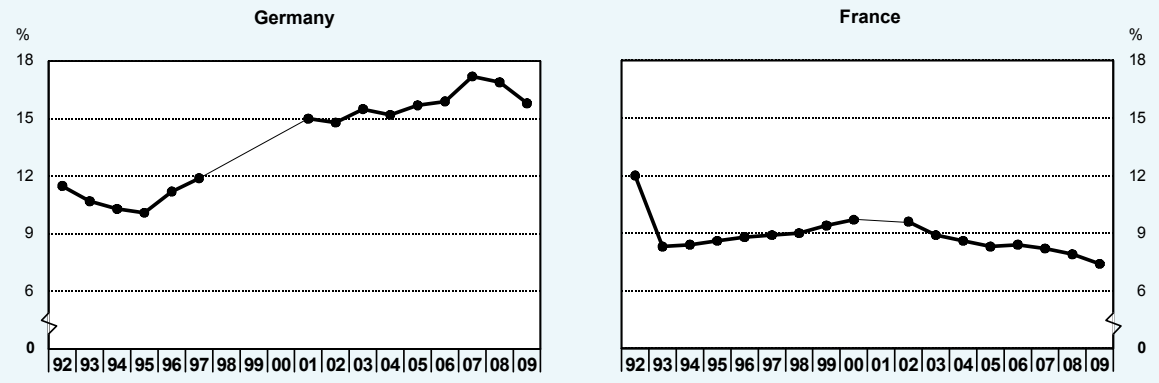
Having made this choice, we are left with at least eleven individual indicators of the third element to be condensed into one. Neither the UNECE task force nor the contributing organisations provide any guidelines for aggregation. For the concept of decent work, a first attempt at aggregation was undertaken by Bonnet et al. (2003), albeit with the aim of identifying basic securities "in society, in the workplace and for individual workers".

145. For the sake of comprehensibility and simplicity, we consider as before a series that we also use to serve as headline indicator for a PCA. Out of the third element of the quality-of-employment framework we choose the percentage of employees working in shifts as an indicator of this dimension. Highly valued personal activities are typically associated with certain hours of the day, so shift work has a negative influence not only on the direct quality of life during work, but also on the quality of life of other personal activities that have to take place "off peak time". As the percentage of German employees working in shifts increased steadily up to 2007, one gets the impression that there is a decrease in quality of life in Germany, while the opposite movement can be seen for France (Chart 17).

Chart 17

Employees working in shift work[1)]

Percent of total employees



1) Source: Labour Force Survey.

Data

146. Given the lack of a convincing alternative aggregation methodology and as a cross-check for this finding, we consider **PCA** to be a fruitful method. We propose starting from the list of indicators provided by Körner et al. (2010) for Germany and by the UNECE Task

Force for France (UNECE Task Force on the Measurement of Quality of Employment, 2010). The chosen data are extracted from the Labour Force Survey for Germany and France.

147. As argued above, we restrict ourselves to the third element of the quality-of-employment framework, "**working hours** and **balancing** work and non-working life". Of the indicators contained therein, the following say something about personal activities. Quality of life can be constrained for economically active persons if an individual wants to work full-time but finds only a part-time job, or works undesirably long hours or at unusual times of the day, with corresponding repercussions on work-life balance. People who involuntarily work part-time are probably dissatisfied with their jobs both because their current job may well not be their preferred choice and because they receive less income than in the preferred full-time job. Consequently, an increase in each of the indicators indicates a less desirable situation. For Germany we use data from the Labour Force Survey from 1993 to 2009. The share of each of the variables increases in the observation period, indicating a deterioration of the state of affairs (Table 12).

Table 12

Personal activities – Variables of the principal component analysis: Germany[1)]

	1993	2009
Share of		
involuntarily part-time workers in the total number of part-time workers	5.6	21.9
employed persons working at night	7.6	8.1
employed persons working on Saturday	21.1	24.8
employed persons working on Sunday	10.4	12.9
shift-workers in the total number of employed	11.5	15.7

1) Source: Labour Force Survey.

Data

148. In our analysis we focus on (dependent) employees only, so that in the case of Germany at most our analysis embraces little more than half of the population. Thus it excludes people who are not economically active, for example retirees, housewives, children as well as students and unemployed persons. Yet these excluded population groups could possibly have specific problems. The unemployed, for example, might be unhappy because they cannot find work and therefore do not allocate their time between leisure and work according to their preferences.

149. In the application to German data, each of the variables has the expected positive weight and the results are robust for different sub-samples. Since in our study an increase in an overall indicator is always intended to indicate an improvement in the quality-of life dimension under scrutiny, we have to multiply this first principal component by minus one (Table 13). Consequently, the overall personal activities indicator yields an explanation of 94% Germany, and the KMO value is 0.77. For Germany the composite personal activities index decreased until the year 2007, which suggests that the situation had worsened. Since then the index has increased, however, indicating an improvement. This pattern replicates that of the interim

headline indicator (Table 13), which makes us more confident to use it. For France, a PCA cannot be performed because of a structural break in the time series of the share of involuntarily part-time workers in the total number of part-time workers and low KMO values that advise against a PCA.

Table 13

Personal activities – Weights of the first principal component: Germany[1)]

Share of	
involuntarily part-time workers in the total number of part-time workers	0.524
employed persons working at night	0.390
employed persons working on Saturday	0.421
employed persons working on Sunday	0.423
shift-workers in the total number of employed	0.466
Kaiser-Meyer-Olkin measure of sampling adequacy	0.770
Eigenvalue of first principal component	5.495
Proportion of Variance explained by first principal component	0.939

1) Calculations based on the Labour Force Survey. Data

Political voice and governance

150. The opportunity to express an opinion in the political sphere – together with the implied absence of political repression – is another source of quality of life. Its consideration as a separate dimension of quality of life is supported by the insights of the **capabilities** approach. To express a political opinion is an essential element of freedom and capabilities. Political voice and governance comprise more than merely counting voter turnouts and party memberships. It covers the functioning of a parliamentary democracy, elements of direct democracy, universal suffrage, civil society organizations, independent media, legislative guarantees, rule of law, and effective enforcement of laws. Many of these elements are **difficult to measure** objectively, and they themselves comprise a plethora of facets.

151. Indicators that combine several sources of information and various aspects of voice and governance have the potential to provide the broad coverage needed to capture this multitude of relevant aspects. However, the respective **composite indicators** have a number of disadvantages that need to be taken into account when interpreting levels and changes. Most researchers concerned with the empirical analysis of this dimension have mainly relied on **expert opinions**. This strategy has clear disadvantages when it comes to assessing the actual and perceived adequacy and fairness of a given institutional set-up. However, **population surveys** which would allow for such an assessment are rarely available regularly and are typically not conducted in a format that would allow for international comparisons. At most**,** surveys about citizen's voice, legislative guarantees and the rule of law can only be considered a future alternative to existing composite indicators.

152. Three existing composite indicators have received particular attention. Freedom House publishes "**Freedom in the World**", in which indicators of "political rights" and "civil liber-

ties" are presented, both on a seven-point ordinal scale. After averaging the two results, a country is classified as "free", "partly free" or "not free" (Freedom House, 2010). In addition, the Centre for Systemic Peace at George Mason University undertakes the **Polity IV** project with an index for "degree of democracy". It is measured on a scale from -10 (full autocracy) to +10 (full democracy). Among the elements included are institutionalized procedures for open, competitive, and deliberative political participation; choosing and replacing chief executives, type of elections; checks and balances on the powers of the chief executive (Marshall and Jaggers, 2007). Third, the World Bank Institute annually publishes six "**Worldwide Governance Indicators**".

Both the Freedom House and the Polity IV indicators have a clear focus on developing and emerging countries. As a consequence, the scaling of the respective indicators does not allow for differentiation between OECD countries like France and Germany, which obtain the maximum possible points in both concepts. Even disregarding aspects of **methodological consistency**, the **political relevance** of these measures for our countries is therefore extremely limited.

153. Among the "Worldwide Governance Indicators" published by the World Bank Institute, the one termed "**voice** and **accountability**" appears to be best suited for our purposes. It is "capturing perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media" (Kaufmann et al., 2009). This indicator is constructed from various data sources that capture either expert opinions or surveys. The data are then used for the estimation of an Unobserved Components Model to extract a minimum-variance estimate of the particular dimension of governance, in our case voice and accountability.

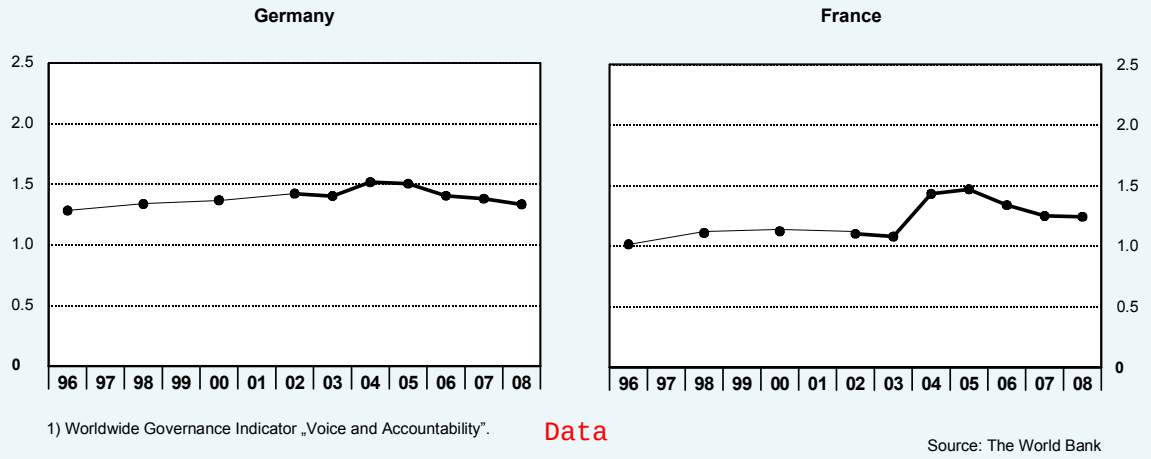
The World Bank Institute presents results in two ways. First, all countries are ranked and percentile ranks are presented, with the highest decile printed in dark green colour. Second, the **government score** ranging from -2.5 to +2.5 is presented, together with its standard error. In the case of France and Germany, in 2008 these countries rank 90.4 and 92.8, and obtain a score of +1.24 and +1.34, respectively, both with a standard error of 0.14. Given its annual publication (since 2002), its international comparability and its broad concept, we regard the government score of the World Bank Institute's "voice and accountability index" as the first choice for an overall indicator of the quality-of-life dimension "political voice and government".

These government scores are presented for France and Germany in Chart 18. Both countries receive values above +1 and hence rank among the **best countries** worldwide. While Germany was in the highest decile for a long time, more recently it has experienced a marginal decrease in its absolute indicator value as well as in its standing relative to other countries. By contrast, according to this specific indicator France has improved from the lower end of the highest quintile to the lower end of the highest decile of countries over the course of the past decade.

Chart 18

Voice and Accountability[1)]

Score ranging from -2.5 to +2.5



1) Worldwide Governance Indicator „Voice and Accountability".

Data

Source: The World Bank

Social connections and relationships

154. Social connections are very important for people's quality of life. They are important because, for example, labour markets are permeated by networks, so that most people tend to find a job through whom they know rather than through what they know (Stiglitz et al., 2009). In addition, social connections bring benefits for health: social isolation rivals smoking as a risk factor for premature death (Berkmann and Glass, 2000). Furthermore, much evidence suggests that social connections are one of the most robust predictors of subjective measures of life satisfaction. Nonetheless, research on social connections is relatively new and statistics are still rudimentary. Moreover, social connections and relationships is the dimension of quality of life that arguably is the least accessible to objective measures. The mere number of family members or persons considered as friends says little about the degree or intensity of social connectedness. It seems advisable to rely on **survey data** with respect to this dimension.

155. In Europe, two particular questions seem to be appropriate candidates for measuring social connections and relationships. First, respondents are asked about their frequency of **spending time** with people at sport, culture or communal organizations. Possible answers are "Weekly", "Once or twice a month", "Only a few times a year", "Not at all". This question, posed in the World Values Survey 1999/2000, has been selected by the European Commission as an objective Well-being indicator for "Social interactions" (European Commission, 2010). The figure recorded is the share of respondents who spend time weekly. The categories "Don't know", "No answer", "Not applicable" and "Missing; Unknown" are not included in the total.

A second possible question is the **ability to ask** relatives, friends or neighbours **for help**. This question was posed for a secondary target variable in the EU-SILC module 2006. In the subtext for the interrogator, it says: "If the respondent has the ability to ask for help from any relative, friend or neighbour. The question is about ability for the respondent to ask for the help whether the respondent has needed it or not, the potential of getting help even if the help

actually has been received or not. Only relatives and friends (or neighbours) who don't live in the same household as the respondent should be considered" (European Commission, 2006).

156. Some critical observers might consider the last question as rather indirect and a little bit abstract. Furthermore, it might be case that someone has many persons who would help him, but at the same time the intensity of this relationship is very low. Nevertheless, one can also argue that someone only asks people who are close friends for help. In our view, the first question about the frequency of spending time with people at sport, culture or communal organizations is more appropriate for illustrating the intensity of social connections and relationships. On the one hand, it is more appropriate because of its more direct and less abstract approach to illustrating social connections and relationships. On the other hand, the frequency of spending time with people seems to be a better indicator of quantity and quality of social connections and relationships than is the ability to seek help. Clearly, this indicator is also related to the dimension of personal activities, especially given the currently chosen interim indicator for that dimension (shift work). However, there is no complete overlap: while in the former dimension quality of life results from the activity (e.g. shift work), in this dimension it results from the people available to interact with.

Thus, we propose to include this question in the annual programme of EU-SILC and its results in our dashboard, as an indicator of social connections and relationships as a dimension of quality of life. Until such data become available, we will not report an indicator for this dimension.

Environmental conditions

157. Environmental conditions affect people's quality of life in several ways. Firstly, they play an important role when it comes to health issues, as the quality of air or water or the noise level have a direct effect not only on bodily, but also on psychological health. Studies show that environmental conditions are the cause of about one quarter of all diseases worldwide (World Health Organisation, 2008). Hence, people benefit strongly from clean water and a healthy nature. Secondly, good environmental conditions are a prerequisite for recreation purposes. Access to natural space as parks, woods or lakes can enhance the variety of potential leisure-time pursuits and is thus one driver of quality of life.

From a long-term perspective, as is argued forcefully in the chapter on sustainability, keeping the environment in a healthy shape is necessary to prevent severe harm to people's life in the future. Strong climate variations bring about, for example, drought/floods or rising sea water levels which pose a risk not only to property and well-being, but also to the provision of basic needs. However, in the context of quality of life, as it is understood here, the short-term view is emphasized.

158. When it comes to the question of how environmental conditions can be captured in a single indicator, different measures come into mind. The choice of the best measure is determined on the one hand by practical aspects, like availability and comparability, and on the other hand by the best approximation to what is to be measured. To start with the latter, a list

of adequate indicators would include the share of population suffering from noise or air pollution, the quality of water, land utilization, distance to natural space (or, in the opposite case, vicinity to industrial plants), and the density of the population. For the level of the environment dimension of quality of life, the climatic situation within a region, as measured, for example, by the sunshine duration could additionally be taken into account.

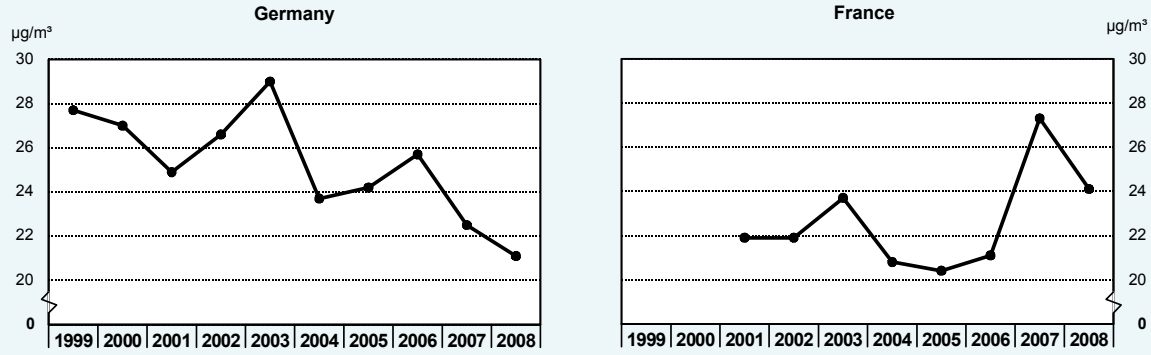
159. Over the past decades, much effort has been spent on measuring environmental conditions in different ways. However, perhaps the most reliable approach, which is to rely on physical aspects rather than on opinion surveys, is used here. A composite indicator which accounts for all facets of quality of life mentioned above would be the first choice. However, this would require an assessment of the relative importance of the individual measures, which is difficult to derive. Using existing composite measures poses the same problem of explaining the different weights.

160. For pragmatic reasons we choose an existing and simple individual indicator as our headline indicator. This indicator measures the urban population exposure to **air pollution** by **particulate matter** of diameter less than 10 micrometers (PM10). This air pollution indicator – which incidentally is also one of the European Commission's sustainable development indicators – shows "the population-weighted annual mean concentration of particulate matter at urban background stations in agglomerations" (i.e. measurement stations that are not directly where the emission takes place), as reported by Eurostat.

The advantage of the chosen indicator is that it not only accounts for air quality alone. Additionally, it indirectly depicts the existence of natural space and is an indicator of the density of traffic or industrial plants, and hence noise pollution. However, the correlation pattern with the other elements needs to be explored in more detail. Data for this indicator often exist even on a daily basis for many developed countries and are highly comparable internationally. For example, there are two measurement stations in the city of Wiesbaden alone. Chart 19 shows that there is a declining trend of air pollution in Germany, whereas the picture is somewhat mixed in France.

Chart 19

Urban population exposure to air pollution by particulate matter[1)]



1) The indicator shows the population weighted annual mean concentration of particulate matter (particulates whose diameter is less than 10 micrometers) at urban background stations in agglomerations.

Data

Source: EU

Personal and economic insecurity

161. Finally, **personal** and **economic insecurity** again calls for a composite indicator to capture at least a sizeable fraction of its facets. These include fear of death, crime, violence, unemployment, illness, and poverty, to name some of the most important ones. While a composite indicator has the advantage of capturing many facets, it has to include valuations of the individual data included in the composite indicator.

162. The Canadian Council on Social Development has come up with a **composite** indicator for personal insecurity that seems to be suitable for our case (Canadian Council on Social Development, 2003). While this organization has two indicators, one for insecurity data and one for the perception of insecurity, our attention will be restricted to the objective indicator. It consists of data from three different areas: economic security, i.e. security about one's job and financial status, health security, i.e. protection against the threats of disease and injury, and physical safety, i.e. feeling safe from violent crime and theft. For each of the three areas, a number of indicators have been selected by professionals.

For **economic security**, the selected indicators are:

- Personal disposable income per capita,
- Poverty gap,
- Long-term unemployment rate,
- Percentage of unemployed receiving employment insurance,
- Average level of social assistance benefits,
- Ratio of total mortgage and consumer debt to total disposable income.

For **health security**, the indicators are:

- Potential years of life lost,
- Incidence of workplace injuries,
- Motor vehicle accident injury rate.

For **physical safety**, the indicators are:

- Violent crimes in relation to 100,000 people,
- Property crimes in relation to 100,000 people.

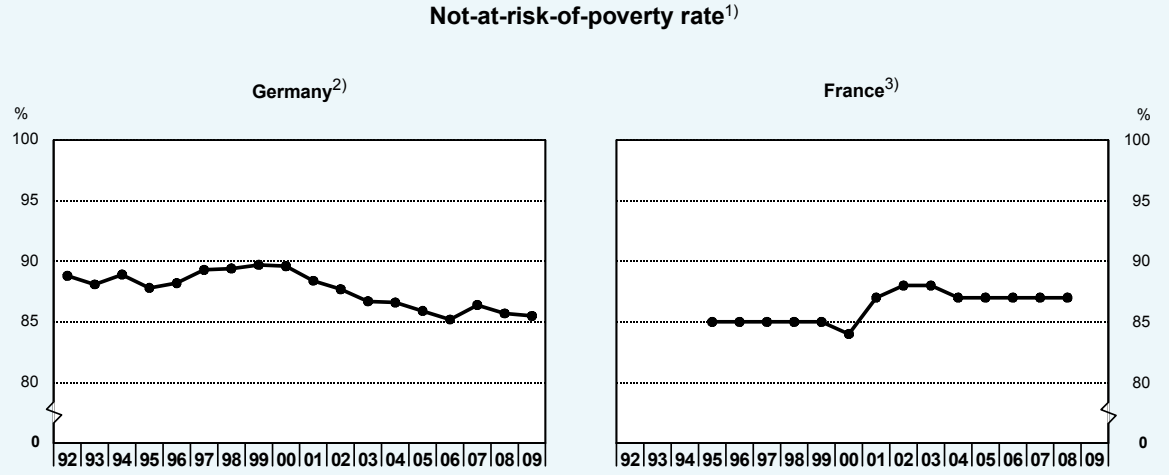
163. Within each of the three areas, all indicators are equally weighted. The weighting of the average indicators of each area is then determined by asking a representative part of the population about the relative importance of the three areas for personal security. The result – economic security accounts for 35 %, health security for 55 % and physical safety for 10 % – is then used to come up with a total personal security index. We regard this procedure as fruitful for adaptation in Europe. To adapt it, first the individual series should be collected. This is already done by Eurostat, the European Central Bank, the Direction centrale de la police judiciaire and the Bundeskriminalamt. Then, one survey question should be added to the regular EU-SILC modules to evaluate the relative importance of the three areas.

164. Given these different facets of insecurity and the current lack of sufficiently long time series, our preferred method PCA is currently not applicable to this dimension. Once this situation improves, PCA should be entertained as well.

For the time being, a single headline indicator should be chosen. Despite serious reservations regarding the use of relative measures of poverty, we propose to rely on the European Commission's headline indicator of social inclusion, i.e. the at-risk-of-poverty rate. It measures the share of persons with an equivalized disposable income below the risk-of-poverty threshold, which is set at 60 % of the national median equivalized disposable income after social transfers. With this indicator we hope to capture the area of economic insecurity at least to some degree. With respect to the remaining two areas, positive correlations between poverty and health insecurity and risk of crime might allow us to temporarily disregard these areas.

Results for the **not-at-risk-of-poverty rate** are presented in Chart 20. While this rate has decreased in Germany in the current decade, it is quite stable in France, and at a slightly higher level. Among other developments, these patterns for Germany might reflect the most recent overhaul of the systems of unemployment support and social security.

Chart 20



1) One minus share of persons with an equivalised disposable income below the risk-of-poverty threshold, which is set at 60 % of the national median equivalised disposable income after social transfers.– 2) Source: SOEP.– 3) Source: EU.

Data

5. Recommendations for future work

165. This study has laid the ground for enhanced **regular reporting** of the state of well-being that comprehensively covers a wide spectrum of facets of human existence. Regarding the results for quality of life, besides a summary of the most recent developments, the intricacies of the matter require that the bald numbers must always be elucidated and interpreted **carefully.** After all, the very nature of the various non-material dimensions of quality of life means that even the best indicators of the state of affairs are only **imperfect proxies** and should be discussed with all due consideration of their potential and their limitations before formulating any recommendation for policy action. Furthermore, we propose to visualize the results in the form of a radar chart which illustrates the developments along the seven dimensions over time and demonstrates the multi-faceted nature of the phenomenon under study.

But one should **never** fall into the seductive trap of constructing an **encompassing quality-of-life indicator** or surface measure, as easy as that might be in terms of calculations.

166. The SSFC report makes five recommendations with respect to quality of life, leaving it to further research to set the adequate priorities between them. First, **measurement** along all but the first dimension should be improved, with particular efforts necessary for social connections and relationships, political voice and governance, and insecurity. Second and third, **inequalities** should be assessed and **interrelations** between the dimensions explored. Fourth, various forms of **aggregation** should be made possible through adequate provision of information. And finally, **subjective measures** of well-being should be surveyed by statistical offices. As they are quite general, the Conseil d'Analyse Économique and the German Council of Economic Experts naturally agree with all of these five uncontroversial recommendations. In our own contribution, we have decided to improve the state of play regarding two areas touched upon by the recommendations with the objective of forming a solid basis for the actual application of the conceptual ideas.

167. The **first contribution** we make is with respect to **aggregation**. The construction of composite indicators is more than a mere technical issue, since it always involves a large range of serious identification assumptions. Our detailed discussion of this matter has led to the formulation of a pragmatic and yet, at least in our own assessment, conceptually sound strategy. While we are adamant that aggregation across the dimensions of quality of life would have to rely on overly strong identification assumptions, aggregation within one dimension might be less controversial. Of the various methods available to aggregate within dimensions, we assess the potential of two of these methods to condense information. Furthermore, our discussion pays considerable attention to the **communication** of the results. In particular, we propose the publication of graphs that visualize the results.

168. The **second contribution** relates to concrete steps towards improving **measurement**. At first glance, measures of the dimensions of quality of life are in abundant supply. Some of its elements – mortality tables, violent crime – even belong to the oldest statistics collected regularly. Yet closer inspection reveals the imperfect state of affairs, as our detailed discussion has documented. Given the intensity of efforts spent by governments and statistical offices on this matter, however, there is ample reason to hope for rapid improvement.

To improve the current state of affairs, one has to survey the existing measures within each dimension and single out the most important **deficiencies**. Major topics in this context are international availability and comparability, both between France and Germany and within Europe. Moreover, the frequency with which the measures are currently calculated is insufficient. A final issue relates to the fact that measures of opportunities instead of achieved functionings might be a particular requirement. Our proposals for each quality of life dimension are outlined below.

169. As the most desirable indicator of the dimension **health** would be the result of a PCA, improvements could be achieved with respect to the timeliness of the underlying indicators.

Furthermore, **morbidity** is so far only covered to a very small extent. Thus, representative annual time series on morbidity data would be highly desirable. With respect to this dimension's envisaged future headline indicator healthy life years, data already exist for Europe as from 1996. Its first element – mortality data – has been in existence for centuries, so the learning curve for quality improvements is rather flat. Its second element – morbidity data – is comparably new, and Eurostat's methodological changes during the past years clearly show that the learning curve is steeper here. Following the transition from the European Community Household Panel (ECHP) to the EU Statistics on Income and Living Conditions (EU-SILC) between 2003 and 2005, the comparability over time is relatively high. Since then, minor changes in the ordering and wording of the survey questions have been made, e.g. in 2008.

Problems remain with respect to cross-country comparability, as people report subjectively on their **disabilities**, and these reports may differ across countries. Two routes for dealing with this issue seem fruitful. First, one could contrast self-perceived illnesses with objective data on the prevalence of illnesses to obtain country-specific correction factors. Second, a more restrictive interpretation of disabilities may lead to higher cross-country comparability of the results. Finally, people living in institutions like old people's homes are currently not covered by EU-SILC.

170. Proposals for the indicator of the dimension **education likewise entail** increasing the timeliness of the indicators underlying the PCA and, even more importantly, engaging in the more intense measurement of education **outputs**. Moreover, with respect to our favourite, yet so far unrealized PIAAC indicator, some minor issues arise related to our specific focus on quality of life. Quality of life increases if people are open to other cultures and other people's views, have learned to express themselves and to discuss, and if people enjoy education. It would be welcome if PIAAC were to keep an eye on these issues as well. Finally, in order to derive representative results for the whole population, it would be desirable to cover all age groups.

171. Given that our chosen indicator of **personal activities** would be the result of a principal component analysis, the underlying vector of time series should be extended to include new or better data sources. Research should be devoted to evaluating the ideal indicators to be collected and selected for the PCA. But as it will take at least ten years of data before a series can be included in the PCA, the current setting will probably not change that soon.

The indicator of the dimension **political voice** and **governance** is an already existing composite headline indicator created by an international organization. The scope for changing its measurement is therefore possibly more restricted. Still, we agree with the SSFC report that **population surveys** are highly valuable for complementing and sometimes replacing experts' views.

The chosen indicator of the dimension **social connections** and **relationships** clearly captures only part of what is intended to be measured. Frequency of spending time with people in cer-

tain circumstances says nothing about the quality and intensity of the connections and relationships. More than elsewhere, the floor is open for debate on a better all-embracing indicator. In the meantime, the proposed headline indicator should be surveyed annually within EU-SILC.

172. Measurement of urban population exposure to air pollution by particulate matter as an indicator of the dimension **environmental conditions has** existed since 1999 for the EU 27. Of course, this indicator again captures only a tiny fraction of environmental conditions, but this fraction is captured well, and it is representative of many more fractions. The quality of this indicator is generally high. Therefore we see no pressing need to improve its measurement.

173. The proposed headline indicator of the dimension **personal** and **economic insecurity has been in** existence since 1998 and is collected by Eurostat. Though timeliness is an issue here, there is no direct need for improvement. The alternative composite index is put together from eleven indicators, all of which have been available for the EU since at least 2002. Room for improvement in this field relates to the aggregation of the three subfields of economic security, health security and physical safety. The population survey to detect the respective weights could be carried out more or less frequently, the sample could be more or less representative. For the time being, the first survey ought to be carried out, and we propose to do this by posing an additional question within an EU-SILC module.

Résumé

174. One does not have to leave the vantage point of economics to realize that life has more to offer than its material aspects. **Non-material** elements of **well-being** play an important role in determining individual fulfillment and satisfaction and societal progress. This chapter has discussed the difficult task of gauging non-material well-being at the individual level and, via the aggregation of individual information, at the level of societies. Moreover, it has provided a first application of the **empirical strategy** emerging from this discussion to the cases of France and Germany, guided by the clear understanding that this analysis is a first step and not an end in itself. In this endeavour, we have made a series of **deliberate choices**, both at the conceptual and the applied level, balancing the desirable with the achievable.

175. Regarding the **conceptual discussion**, we strongly advocate what we have termed a **bottom-up approach**. We could have started our search for a better grasp of the state of non-material well-being from interview information on individual "happiness", but fundamental questions of measurability and the risk that such inherently imperfectly defined measures of human satisfaction could too easily be manipulated into showing politically desirable results have prevented us from embracing this approach. Instead, our advice would be to condense the ample information on diverse elements of non-material well-being as much as possible so as to make the information digestible by its recipients while simultaneously retaining as much of its complexity as necessary to reflect its variegated nature.

Our concrete empirical strategy starts from the definition of a range of **dimensions** which should not be aggregated any further in order to adequately capture life's complexity. In our application, we have been guided by the SSFC report into choosing seven dimensions, some of which pertain to individuals themselves, such as health and education, while others describe the societal and physical context experienced by individuals, such as social connections and relationships and environmental conditions. The strategy then proceeds from dimension to dimension, one at a time, and identifies for each of them a series of **individual indicators** which capture its facets as comprehensively as possible. Finally, for each dimension separately, we select one **headline indicator** out of this reservoir to represent the dimension as well as possible.

Whenever feasible, we engage in a procedure of statistical complexity reduction in order to cross-check our selection of headline indicators. Most importantly, throughout our analysis we have worked under the **constraint** that the indicators chosen need to be regularly available in order to facilitate a perpetuation of this report in future years.

176. The application of this strategy to two countries, France and Germany, has uncovered a set of results that are plausible in that they paint a **mixed portrait** of **societal progress** over the last decade. In particular, progress in terms of health, education (with some reservations), and environmental conditions appears to be highly congruent with the steady growth experienced in material well-being. And yet, while they are admittedly difficult to capture, the recent developments in other dimensions of non-material well-being, such as personal activities and personal insecurity, indicate that societal progress has not been achieved unequivocally across all relevant dimensions.

References

Berkmann, L. F. and T. Glass (2000) *Social Integration, Social Networks, Social Support, and Health*, Social Epidemiology.

Bonnet, F., J. B. Figueiredo and G. Standing (2003) *A Family of Decent Work Indexes*, International Labour Review, 142 (2).

Bynner, J. and S. Parsons (1997) *It Doesn't Get Any Better. The Impact of Poor Basic Skills on the Lives of 37 Year Olds,* London: The Basic Skills Agency.

Canadian Council on Social Development (2003) *Personal Security Index. A Reflection of how Canadians Feel. Five Years Later.*

Chernyshev, I. (2009) *Linkages between the Decent Work Concept and the Quality of Employment Frameworks*, Presentation, Geneva 2009-16-10, UNECE/ILO/Eurostat Meeting on the Measurement of Quality of Employment.

European Commission (2006) *Description of SILC User Database Secondary Target Variables: Module 2006 Social Participation*.

--- (2010) *Well Being Indicators. Feasibility Study. Data Presentation by Countries*.

Ferrer-i-Carbonell, A. and P. Frijters (2004) *How Important Is Methodology for the Estimates of the Determinants of Happiness?*, Economic Journal, 114 (497), 641 - 659.

Freedom House (2010) *Freedom in the World 2010. Erosion of Freedom Intensifies*.

Frey, B. S. and A. Stutzer (2001) *Happiness and Economics: How the Economy and Institutions Affect Human Well-being*, University Presses of California.

Frijters, P., J. P. Haisken-DeNew and M. A. Shields (2004) *Investigating the Patterns and Determinants of Life Satisfaction in Germany Following Reunification*, Journal of Human Resources, 39 (3), 649 - 674.

Giovannini, E., J. Hall, A. Morrone and G. Ranuzzi (2009) *A Framework to Measure the Progress of Societies*, Draft OECD Working Paper, September 2009.

Helliwell, J. F. (2003) *How's Life? Combining Individual and National Variables to Explain Subjective Well-Being*, Economic Modelling, 20 (2), 331 - 360.

Helliwell, J. F. and C. P. Barrington-Leigh (2010) *Measuring and Understanding Subjective Well-Being*, http://www.nber.org/papers/w15887.bib (accessed 2010-04-14).

International Labour Office (2008) *Nineteenth Item on the Agenda. Report of the Director General: Third Supplementary Report. Tripartite Meeting of Experts on the Measurement of Decent Work*.

Jolliffe, I. (2002) *Principal Component Analysis*, 2. ed., New York: Springer.

Kaiser, H. (1970) *A Second Generation Little Jiffy*, Psychometrika, 35 (4), 401 - 415.

--- (1974) *An Index of Factorial Simplicity*, Psychometrika, 39, 31 - 36.

Kaufmann, D., A. Kraay and M. Mastruzzi (2009) *Governance Matters VIII Aggregate and Individual Governance Indicators 1996 – 2008*, World Bank Policy Research Working Paper 4978.

Kirsch, I. (2001) *The International Adult Literacy Survey (IALS): Understanding What Was Measured,* ETS Research Report, December 2001, RR-01-25 , Statistics & Research Division, Princeton, NJ.

Körner, T., K. Puch and C. Wingerter (2010) *Wie wir Arbeiten. Indikatorenbericht zur Qualität der Erwerbstätigkeit*, Wiesbaden: Statistisches Bundesamt.

Layard, P. (2005) *Happiness: Lessons from a New Science*, New York: Penguin Press.

Marshall, M. and K. Jaggers (2007) *Polity IV Project. Political Regime Characteristics and Transitions, 1800 – 2007. Dataset Users' Manual*.

Nussbaum, M. C. (2000) *Women and Human Development: The Capabilities Approach*, Cambridge: Cambridge University Press.

OECD (2008) *Handbook on Constructing Composite Indicators*, Paris: OECD Publications.

Sen, A. (1999) *Commodities and Capabilities*, Oxford University Press.

Stiglitz, J. E., A. Sen and J. Fitoussi (2009) *Report by the Commission on the Measurement of Economic Performance and Social Progress*.

UNECE Task Force on the Measurement of Quality of Employment (2010) *Measuring Quality of Employment*, Geneva: United Nations.

World Health Organisation (2008) *The Global Burden of Disease: 2004 update*.

CHAPTER IV

Sustainability

Sustainability

177. While previous chapters were concerned with the measurement of current economic performance, material well-being and quality of life, this chapter takes a different perspective and addresses the issue of **sustainability**. Essentially, we are concerned with the question "whether we can hope to see current levels of well-being at least maintained for future periods or future generations", which constitutes the core of the third pillar of the SSFC Report. However, in contrast to the SSFC Report, which mainly discusses sustainability with respect to its environmental facets, we take a **broader approach** and include the sustainability of fiscal policies and private sector activities.

Assessing the sustainability of specific economic activities and policies requires widening the perspective in a way which is hardly innocuous. Whenever we discuss the issue of sustainability, we move from recording the actual state of affairs to constructing projections into the future. This is **not the same as** constructing **predictions** of future developments, since the formation of predictions involves an assessment of the likelihood of future events. Instead, the discussion of sustainability is concerned with the consequences of a **persistent prolongation** of current activities and decisions into the future. In effect, statements about sustainability are **what-if statements**, documenting the possible consequences of given paths of action.

Monitoring the sustainability of current policies enables citizens to realize that policies implemented today may have a drastic impact on the well-being of **future generations** or even of the current generation a few years down the road. For instance, by deciding to accumulate public debt, a society can seriously restrict the consumption possibilities of future generations. Predictions, by contrast, necessarily take into account that decision makers might tend to reverse their course of action under the impression of its emerging negative consequences. Since their construction requires numerous behavioural assumptions, predictions are never unequivocal. Thus, predictions undoubtedly should not be the concern of regular statistical reporting.

1. Conceptual issues: dimensions of sustainability

178. Before the issue of sustainability can be meaningfully discussed, it needs to be clearly defined. Various formulations have been proposed in the literature. On the one hand, sustainable development might be described as a situation in which the **current level of well-being** can at least be **maintained** for future generations. On the other hand, sustainability might be viewed as a situation in which future generations have the same **freedom of choice** as any generation before them.

While the second definition is very appealing from the perspective of economic theory, its relevance in practice is limited by the fact that the current generation lacks **information** with which to evaluate with sufficient precision the opportunity set of future generations. For example, it is already difficult enough to project the stock of a given natural resource at current extraction levels, but it might arguably be impossible to take into account the importance of this particular resource for the opportunity set of future generations. After all, this would re-

quire **restrictive assumptions** about technological change and innovation as well as about the preferences of future generations.

In contrast, the first definition is less ambitious, since it merely **extrapolates** the current situation under the implicit assumption that the sources of well-being are the same now and in the future. This is the definition which we will – in line with the SSFC Report – predominantly apply ourselves. While this approach is often more viable in practical work, it should be borne in mind that it tends to understate the sustainability of affairs, since it does not take into account the way in which preferences or technologies might change. It is nevertheless the appropriate choice for our purposes, as it is our **pragmatic** objective to provide a set of indicators which **signal** to the public whether or not continuing along a current path of action might endanger future well-being.

179. The standard conceptualization of sustainability encompasses **three** essential **dimensions**, namely social sustainability, economic sustainability and environmental sustainability (in the spirit of Harris et al., 2001).

- A **socially** sustainable system must achieve fairness in distribution and opportunity, adequate provision of social services including health and education, gender equity, and political accountability and participation.

- An **economically** sustainable system must be able to produce goods and services on a continuing basis, to maintain manageable levels of government and external debt, and to avoid extreme imbalances between different sectors.

- An **environmentally** sustainable system must maintain a stable resource base, avoiding over-exploitation of renewable resource systems or environmental sink functions, and depleting non-renewable resources only to the extent that investment is made in adequate substitutes. This includes maintenance of biodiversity, atmospheric stability and other ecosystem functions not ordinarily classified as economic resources.

Satisfying these three requirements simultaneously is essential for achieving sustainability and the production of well-being. Moreover, since these three goals are highly interrelated, they can only be meaningfully discussed under a multidisciplinary approach. Finally, in a highly interconnected world, their analysis demands a decisively **international** perspective. This is obvious, especially, for environmental issues, as pollutants do not recognize borders. But it holds more generally, since the forward-looking perspective of sustainability discussions incorporates the interaction between many economic agents, and since national policy decisions are typically made in an international context.

180. The first sustainability dimension, **social sustainability**, emphasizes various facets of well-being and quality of life that were already discussed in chapters II and III. Issues central to the cohesion of society, such as the distribution of income or access to high-quality work, have already been reviewed in the **second chapter** on monitoring current material well-being,

leading to the concrete proposal of two indicators for our dashboard. When assessing the question of social sustainability, these facets have to be embedded in an intertemporal and intergenerational perspective. Moreover, as everyday life has a distinct **regional dimension**, one has to analyze social cohesion within and across individual communities. If this analysis reveals that centrifugal forces tend to get the upper hand, cooperation – which is among the major preconditions of social sustainability and societal well-being – is endangered. Yet, this is hardly an issue for the regular reporting conducted by national statistical offices.

Moreover, many indicators that monitor aspects of the current state of quality of life, for example educational participation, measures of income and wealth distribution, access to the labour market, health or political participation, are well suited to monitoring forces that may put social cohesion at risk. These indicators were the topic of the **third chapter**. As the discussion of sustainability generally addresses the question of whether current paths of action, if continued persistently, might have dramatic negative implications, there will necessarily be a **high congruence** between this current assessment of social conditions and their extrapolation into the future. Thus, separate treatment of social sustainability is unlikely to add meaningful information to the indicators of current material well-being and quality of life which are provided within the framework of the second and third chapters.

Moving beyond a straightforward extrapolation will be more than difficult. While it is certainly true that equal opportunities across generations or instances of social sclerosis or social immobility cannot be monitored fully with these indicators, any conclusion regarding social sustainability across generations requires **extremely restrictive** identification assumptions. This arguably exceeds the mandate of any regular statistical reporting imaginable. Specifically, to measure chances and opportunities across generations would require much more information than is currently available. Comfortingly, though, the indicators identified in Chapter III facilitate the identification of any dangerous lack of social capital. One graphic example of intergenerational unfairness is reflected in indicators that take into account the social stratification of educational attainment or performance. As intergenerational persistence and low social mobility tend to be closely related, current information will be highly indicative of long-term consequences.

181. We conclude from this discussion that it is highly advisable to narrow the discussion of monitoring sustainability to **economic** sustainability, comprising macroeconomic sustainability and financial sustainability, and **environmental** sustainability. Macroeconomic and financial sustainability will be discussed in sections 2 and 3 of this chapter, respectively, while section 4 addresses environmental sustainability. Everything that one can convincingly say about social sustainability is already addressed by the indicators developed in the second and the third chapter. In our quest for suitable indicators, we strive for the presentation of indicators following the **principles** of parsimony and practicability, without compromising economic content. In line with the two preceding chapters, we therefore evaluate existing indicators of both economic and environmental sustainability and discuss the status quo of their measurement. We also indicate how the measurement and calculation of existing indicators could be enhanced.

182. We first turn to the discussion of the various facets of economic sustainability. To organize our thoughts, we partition the issue of economic sustainability into three areas.

- **Macroeconomic sustainability** can be separated into growth sustainability, on the one hand, and external and fiscal sustainability, on the other. **Growth sustainability** is the most obvious dimension of economic sustainability. We consider growth to be sustainable if a sufficient part of the creation of wealth is allocated to investment. This investment can be either material, in machinery or infrastructure, or immaterial, in knowledge or skills. For instance, since R&D efforts do seem to be particularly important for future growth, strengthening R&D investment was an important target of the European Union's Lisbon Agenda and has also become a headline target of the EU 2020 strategy.

- **External** and **fiscal sustainability** are related to intertemporal budget constraints of the public and private sector. External sustainability is concerned with the sum of public and private sector balance. Excessive public and private sector deficits, implying an unsustainable external position, can also lead to short and medium-term consequences when current account imbalances are suddenly unwound. Fiscal sustainability refers to the fact that governments can shift the financial burden of current expenditures onto future generations through the intertemporal budget constraint. This issue is closely linked to concerns of intergenerational equity due to its long-term perspective.

- Private sector **financial sustainability** is a concern that is predominantly focused on the medium term, as the build-up and unravelling of financial imbalances ("bubbles") often occurs over just one business cycle. Again, there is a link to long-term concerns, however, as financial crises usually increase public debt levels and thus the burden to be carried by future generations.

2. Macroeconomic sustainability

183. Without any doubt, the global financial crisis that unfolded in 2007 was the consequence of unsustainable economic developments. One of the main lessons of the crisis is that periods of strong growth in GDP can partly reflect the build-up of **imbalances** that are likely to result in **sharp contractions**. And the consequences of these contractions can be quite severe. As governments had to step in to support the financial system as well as domestic demand through stimulus programmes, the state of public finances has deteriorated further. Under this impression, issues of **economic sustainability** have recently taken centre stage in the public debate and in the political discourse and, correspondingly, indicators assessing the sustainability of economic development need to be included in any dashboard that aims at providing a reliable picture of the state of our societies.

Another lesson of the crisis is that through globalization most economies have become so interrelated that no country can be entirely protected from events that happen elsewhere, even if this country does not bear any responsibility for them. Therefore, especially in times of crisis, a certain amount of **international cooperation** might be in the interest of all countries. This is arguably already the case whenever an unsustainable state of affairs is building up,

since then it might still possible to prevent the crisis through corrective action. Yet, international cooperation is often quite difficult to implement given the heterogeneity of national goals. This is particularly the case before a crisis has actually materialized, since without appropriate indicators observers might not be able to realize that an **unsustainable situation** is in fact **emerging**. The inclusion of indicators of economic sustainability in our dashboard is therefore intended to provide the basis for a better informed discussion of international macroeconomic affairs.

In general, this aim should arguably be easier to achieve within a well-defined geographic area like **Europe** than on a wider geographic scale. And many observers would agree that the need for a renewed focus on issues of economic sustainability is especially obvious in the case of the euro area, where recent developments have shown how an increase in asset prices combined with large private and public sector deficits can lead to an unsustainable situation in one country which in the end spills over to other countries.

184. Indicators of environmental sustainability usually emphasize the cost of current modes of behaviour for future generations. A reference to future generations and **long-term** developments is indeed an indispensable element of any meaningful discussion of sustainability. Economists, however, are also concerned about facets of economic sustainability and, thus, use the term "sustainable" in a **medium-term** context that has received much prominence in the context of the financial and economic crisis. According to this perspective, periods of high growth can be considered unsustainable if they are based on changes in the **balance sheets** of households, non-financial and financial firms, the government or the economy as a whole that make sharp and painful adjustments in the future very likely. Specifically, the indebtedness of different segments of an economy (in the case of financial and fiscal imbalances defined more narrowly) or of different regions of the world (in the case of unsustainable current account deficits) can reach levels at which the unwinding of imbalances almost inevitably will take the form of socially costly crises.

In our view, a discussion of sustainability must keep an eye on **both** the very long-term and the medium-to-long-term horizon. We see one of our major contributions as taking the lessons of the crisis to heart by complementing environmental sustainability indicators with information on the economic sustainability of current patterns of growth. This approach is motivated by an additional argument. The well-being of future generations is **closely related** to what happens in the medium term. In particular, the high fiscal cost of financial and balance-of-payment crises usually implies that the leeway for future fiscal policy is substantially reduced. In that case the scope for investment in the well-being of future generations is constrained, for example in the area of fostering the development of environmental technologies or human capital accumulation.

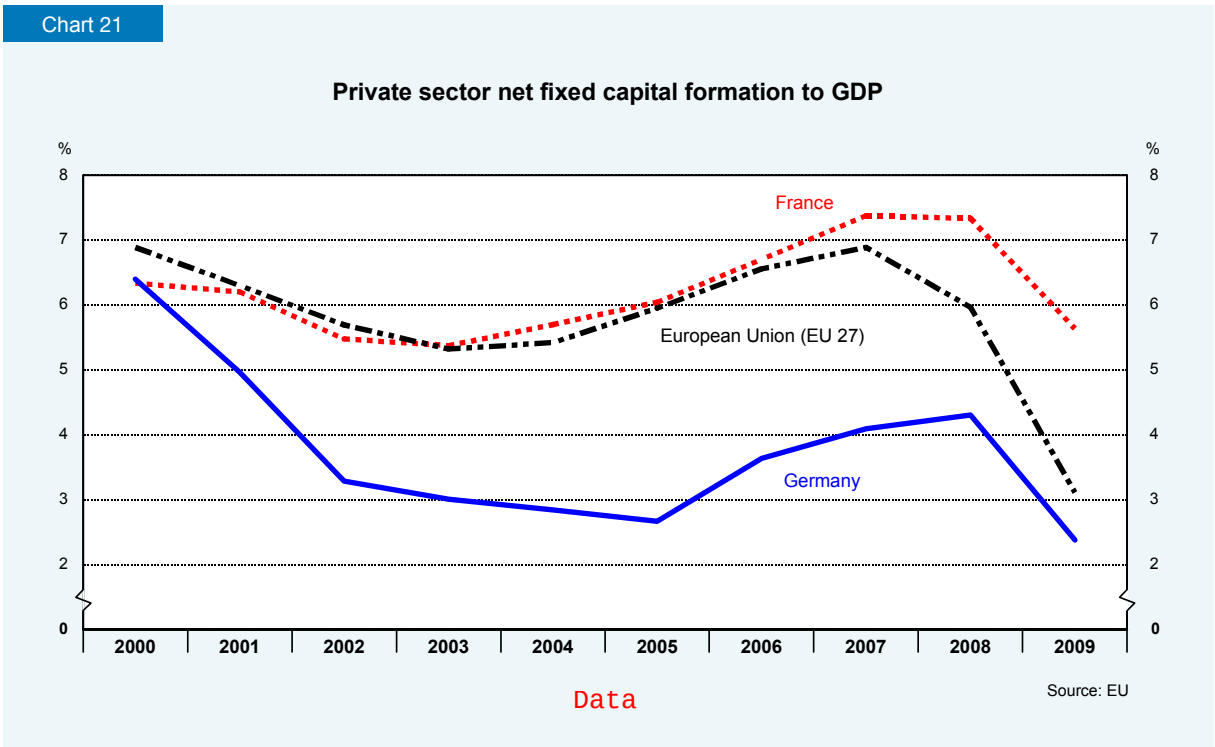
Growth sustainability

185. Well-being depends to a large extent on consumption possibilities and hence on the capacity of producing goods and services. Often it is helpful to think about the production capacity of an economy in terms of a standard growth accounting framework. Within this

framework, the production capacity – or potential output – is determined by three factors: the human capital stock, the physical capital stock and total factor productivity, which determines how efficiently capital and labour are combined in production. To monitor whether an economy is on a path towards a long-run expansion of its productive capacity we therefore need an indicator of each of the three drivers of potential output. Two important determinants of the human capital stock are the quality of the education system – one of the major issues addressed in Chapter III – and the size of the labour force. The labour force participation rate, one of its decisive ingredients, was already introduced in Chapter II. In this chapter we can therefore concentrate on developing indicators for the expansion of the physical capital stock and the improvement of total factor productivity.

186. The development of the **physical capital stock** depends on the amount of gross investment and on depreciation. While gross investment adds to the capital stock, depreciation of existing capital exerts an offsetting effect. This is why the level of **net fixed capital formation** of the private sector (investment net of depreciation) is arguably the most important variable to monitor in this context (see Chart 21 for recent developments of net capital formation in France and Germany). Consequently, in order to emphasize the importance of capital accumulation for economic growth, we conclude that the ratio of net investment (capital formation) to GDP should be added to the dashboard.

Chart 21



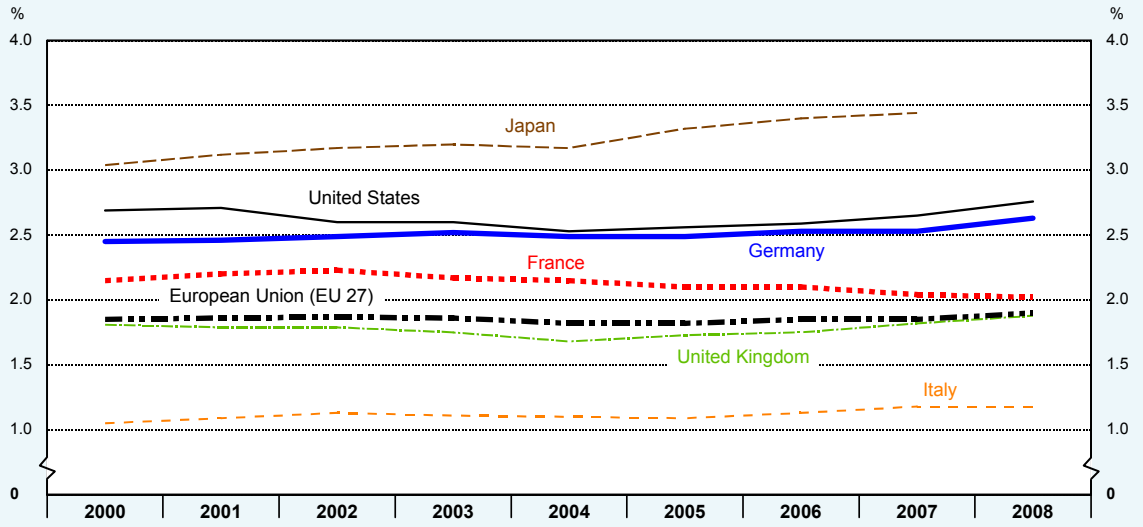
187. Besides investment in the physical capital stock, investment in education, the accumulation of knowledge, innovation and research efforts are among the major determinants of total productivity growth in the long run, as has been established by theoretical and empirical research (see for instance Romer, 1990; Griliches and Lichtenberg, 1982; Griliches, 1986; Howitt, 2000; Jones, 2002). R&D expenditures are the amount of resources allocated to basic, applied and experimental research (irrespective of the source of funds) undertaken by organi-

zations (business enterprises, higher education and research institutions). And business enterprise R&D expenditures are designed to improve the economic performance of firms as well as their productive efficiency and competitiveness. Thus, a measure of **research and development (R&D)** investment of an economy relative to its GDP might serve as a reliable predictor of its future overall productivity, and of expected trends in science, technology and innovation.

Chart 22 shows that the average R&D investment in the EU 27 lay slightly below 2 per cent of GDP between 2000 and 2008. R&D efforts could also be measured by a number of other indicators like the number of patents awarded to each country, or the breakdown of R&D into its public and its private components, or its allocation across sectors. However, since we aim to create a parsimonious dashboard with only a few indicators, we suggest including only the **ratio of R&D investment to GDP**. Not surprisingly, this indicator has also been chosen to measure progress towards the goals set out in the new Europe 2020 strategy (European Commission, 2010).

Chart 22

Research and development (R&D) investment to GDP[1)]



1) Gross domestic expenditure on R&D as a percentage of GDP. „Research and experimental development (R&D) comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society and the use of this stock of knowledge to devise new applications" (Frascati Manual, 2002 edition, § 63). R&D is an activity where there are significant transfers of resources between units, organisations and sectors and it is important to trace the flow of R&D funds.

Source: EU

Data

External sustainability

188. By definition, the **current account balance** equals the difference between savings and investment in an economy in a given year. Savings can be divided into public and private sector savings. A lack of external sustainability, i.e. an unsustainable current account, can hence be caused by unsustainable public or private sector indebtedness. The current account balance is also equal to a country's **net lending** or **net borrowing**. A country that runs a current account surplus exports financial savings (reflected, ceteris paribus, in a negative financial account) and a country that runs a current account deficit imports financial savings (reflected,

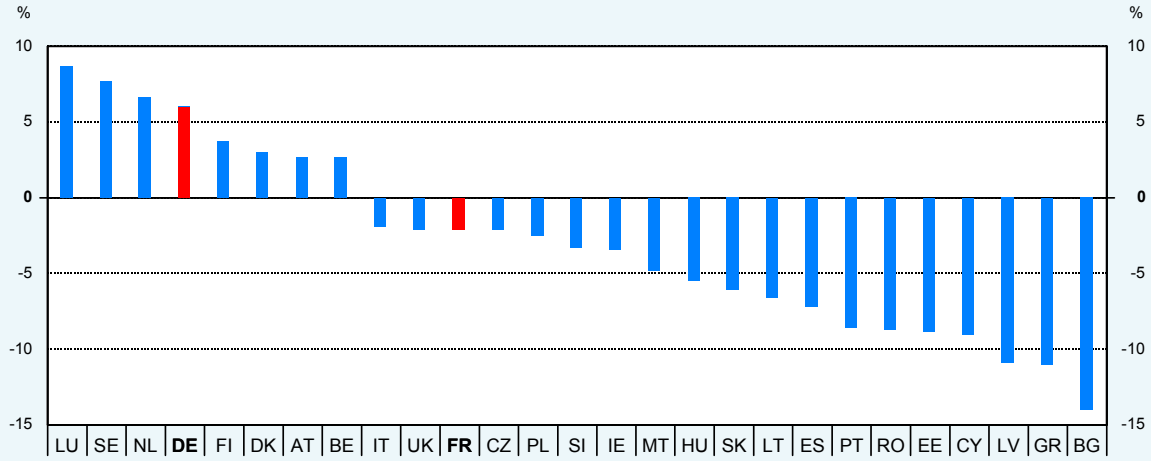
ceteris paribus, in a positive financial account). A current account deficit may be associated with a deficit in the public sector, a deficit in the private sector balance, or both. In all three cases, the current account deficit is financed by **capital inflows** from abroad. This can potentially become problematic when capital inflows are financing unsustainable bubbles or unproductive government expenditures and are suddenly reversed.

189. In the last decade, the **world economy** was indeed characterized by the build-up of large current account surpluses and deficits. In the EU 27 some countries ran current account surpluses of more than five per cent of GDP over the period from 2004 until 2009. Some of the deficit countries showed current account deficits of a similar size (Chart 23). As recent experience has shown, large current account deficits often mirrored unsustainable public sector or private sector imbalances. Closely monitoring the current account can therefore help identify unsustainable developments.

Chart 23

EU 27 countries: current account balances

Percent of GDP: average 2004 – 2009



1) Considered countries: LU-Luxembourg, SE-Sweden, NL-Netherlands, DE-Germany, FI-Finland, DK-Denmark, AT-Austria, BE-Belgium, IT-Italy, UK-United Kingdom, FR-France, CZ-Czech Republic, PL-Poland, SI-Slovenia, IE-Ireland, MT-Malta, HU-Hungary, SK-Slovakia, LT-Lithuania, ES-Spain, PT-Portugal, RO-Romania, EE-Estonia, CY-Cyprus, LV-Latvia, GR-Greece and BG-Bulgaria.

Source: EU

Data

190. For **developing countries**, a current account deficit may be the result of foreign investment. If this investment is profitable, the external debt will later be reimbursed, rendering the accumulation of deficits unproblematic. For **developed countries**, a sustained deficit or surplus in the current account is more questionable. For demographic reasons a country may follow a strategy of accumulating foreign assets in order to prepare for bearing the costs associated with an **ageing population** in years to come. This is quite reasonable. But a chronic current account deficit may also be the result of a lack of competitiveness of its private and public sector. Such a case would indeed be a reason for concern.

It is therefore important to monitor the potential sources of a current account deficit. In particular, it is important to know whether its source is an unsustainable public sector deficit,

originating from excessive public spending, or a continuing private sector deficit, due to unproductive investment or insufficient saving. Indicators of both public and private sector sustainability problems will be included in the dashboard. The discussion of private sector imbalances will be covered in the section on financial sustainability later in the chapter. We first turn to the issue of fiscal sustainability.

Fiscal sustainability

191. Fiscal sustainability has an important impact on the well-being of future generations. By running **unsustainable fiscal policies** over several years, policy makers can impose considerable fiscal burdens on future generations, forcing them to raise taxes or cut spending in order to pay for the consumption of previous generations. Evaluating the sustainability of fiscal policies therefore mainly requires taking into account long-term developments that are best captured by the government's **intertemporal budget constraint**. But unsustainable fiscal positions can also have painful short and medium-term consequences that are a major motivation for the need to closely monitor the present level of public debt, too.

Over the short and medium term, public debt levels affect well-being mainly through two channels. First, high levels of public debt may **crowd out** private investment, thus lowering potential growth in the medium term. Reinhart and Rogoff (2010a), for instance, analyze the correlation between government debt and GDP growth in a sample of 20 countries between 1946 and 2009. They argue that this correlation is weak for debt-to-GDP ratios below a threshold of 90 per cent. Countries above that threshold experienced median growth rates that were approximately 1 percentage point lower, however. Second, when an economy is hit by a sizeable negative shock, such as a financial crisis or a collapse in world trade, governments need room for fiscal manoeuvre in order to be able to respond in a **countercyclical** fashion. This became particularly apparent in the current crisis, when high government deficits and debt ratios already prevented some economies from enacting fiscal stimuli (Horton and Ivanova, 2009).

192. Ideally, rules aimed at ensuring the sustainability of public finances should not prevent governments from contributing to economic growth through public investment. According to the Golden Rule of fiscal policy, the public sector should borrow solely in order to invest and not to fund current spending. In practical applications, however, the central question is how to define investment. As argued in Saint-Etienne (2004) and, in a similar vein, by the Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2007) the correct concept of investment to be considered is **net investment**, since only net investment is able to create new wealth. In Saint-Etienne (2004) it is shown that on average, to take an example, net public investment in the European Union is close to 1 per cent of GDP.

193. In Europe, fiscal sustainability has become a cornerstone of budgetary policies. In the Treaty on the Functioning of the **European Union** member states committed themselves to ensure sustainable public finances (Articles 119 and 120) and to avoid running excessive government deficits (Article 126). These treaties define a sustainable fiscal policy with the help of two reference values for the current government deficit and the stock of government debt. An

annex to the Stability and Growth Pact (SGP), which specified the requirements emerging from these agreements and was adopted in 1997, requires that these reference values should not exceed 3 per cent and 60 per cent of GDP, respectively.

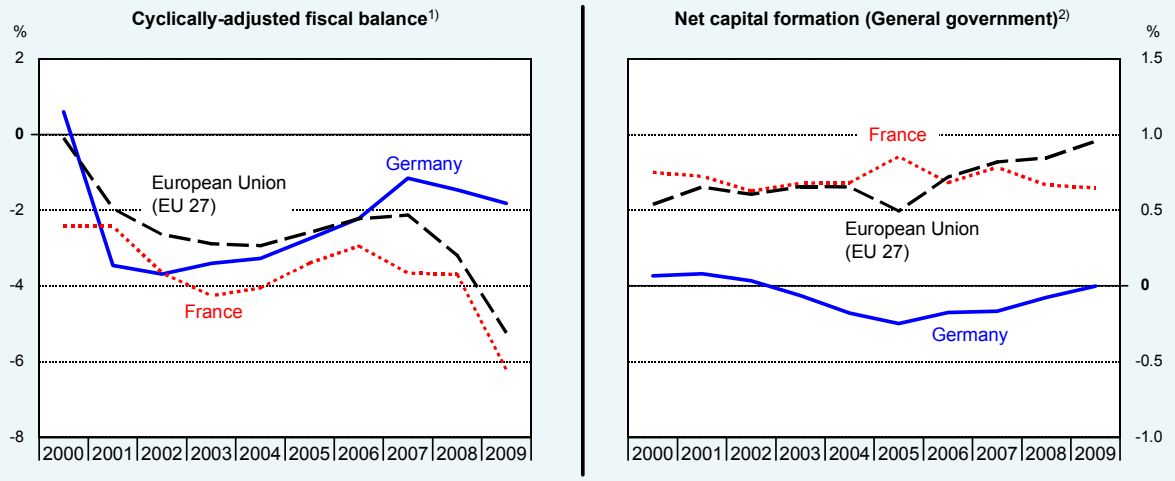
When the SGP was adopted in 1997, it was decided that the ceilings of 3 per cent of GDP for the public sector deficit and 60 per cent of GDP for the public debt were to be maintained, but that these ceilings would be supplemented by the following rules, with "**structural deficits**" measuring the fiscal stance, while correcting for business cycle and temporary effects:

- A country with low public debt and strong growth should aim for an average annual structural deficit smaller than 1 per cent of GDP over an economic cycle,

- a country with high public debt and low growth rates should aim for a positive average annual structural balance over an economic cycle.

Chart 24

Cyclically-adjusted fiscal balance and net capital formation



1) In relation to potential output in percent.– 2) In Relation to nominal gross domestic product in percent.

Source: EU

Data

In what follows, however, we focus on the cyclically-adjusted balance. The only difference with regard to the structural balance is that the cyclically-adjusted balance does not take into account temporary effects such as windfall revenues. An estimate of the cyclically-adjusted balance is easily available across countries in the EU and frequently reported by the European Commission. Because of its importance for the evolution of the public debt-to-GDP ratio over the medium and long term, we suggest including the **cyclically-adjusted balance** in the dashboard as a first indicator of the sustainability of public finances. According to the Golden Rule of Public Finance, the cyclically-adjusted fiscal deficit has to be evaluated in relation to **public net investment**, which it should not exceed. Yet in reality, as Chart 24 documents, in every year since 2001 the cyclically-adjusted deficit exceeded public net investment in both

Germany and France, indicating a problem with the sustainability of public finances in these countries.

194. Since the cyclically-adjusted balance does not account for **implicit liabilities** of the government which might, for instance, arise as the consequence of a public pension scheme, it does not capture the full impact of current fiscal policies on future generations. Thus, to facilitate an assessment of whether current policies allow future generations to maintain the current level of well-being, the cyclically-adjusted balance should be augmented by a second indicator that accounts for all those future government revenues and expenditures that are implied by current policies.

This more comprehensive indicator of the sustainability of public finances must be based on a government's **intertemporal budget constraint**. Its principal elements are the complete paths of government revenues and expenditures into the indefinite future. Since future paths have not yet materialized, the comprehensive assessment of the fiscal stance would therefore need to be based on **projections** of all future explicit and implicit obligations arising from current policies. This is a far-from-easy task. Furthermore, any attempt to compare government expenditure and revenue across different periods requires looking at the present value of projected revenue and expenditure streams. Over the long run, the present value of government revenues must equal the present value of public expenditure in order for the intertemporal budget constraint to hold.

From an ex-post perspective, the intertemporal budget identity is always fulfilled. However, from an ex-ante perspective, it might be violated if the net present value of government projected expenditures exceeds the present value of future revenues, given current policies. In this case there is a **fiscal sustainability gap**, indicating the present value of the fiscal burden faced by future generations. Such a sustainability gap suggests that, sooner or later, the government would have to reduce its budget deficit, either by **cutting spending** or by **increasing taxes**. Failure to adjust fiscal policies enough to close it would allow the ratio of government debt to GDP to get out of hand. Of course, the larger the sustainability gap, the more drastic are the necessary future adjustments of fiscal policy implied by current policies.

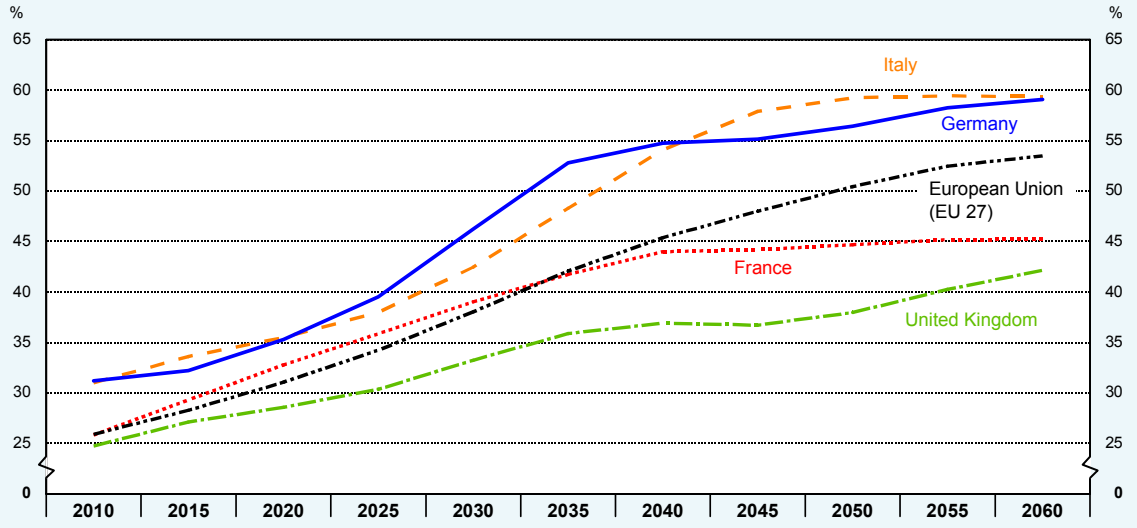
Countries experiencing a drastic **ageing** of the population face more serious policy adjustments in order to balance their intertemporal budget constraint. This problem is particularly severe in Europe, where the average old-age dependency ratio is expected to rise from below 30 per cent in 2010 to almost 55 per cent in 2060 (Chart 25).

195. This suggests that there are a number of ways to assess the severity of the fiscal or social adjustments necessary to close a given sustainability gap. Although all indicators are conceptually equivalent, some of them are more easily interpretable than others. First, there is the possibility to calculate a **sustainable tax rate** (see for instance Blanchard et al., 1990). Given actual forecasts of expenditure and revenue and considering the initial level of debt, the sustainable tax rate would, if implemented now and remaining constant forever, just balance the

intertemporal budget constraint. An indicator of the sustainability of public finances would then report the difference between the actual and the sustainable tax rate.

Chart 25

Projected old-age dependency ratio in the EU[1)]



1) Population over 64 years as percentage of population between 15 and 64 years.

Data

Source: EU

From a purely mathematical point of view, increases in the tax rate and reductions of planned expenditures are equally well-suited to close a sustainability gap. From an economic point of view, expenditure cuts will, however, elicit different behavioural responses than tax increases. Elimination of a sustainability gap should therefore in most cases include adjustments on the expenditure side of the public budget. It therefore makes more sense to represent the size of the sustainability gap as the **permanent reduction** in the public deficit necessary to balance the intertemporal budget constraint, highlighting the fact that a combination of measures on the revenue and spending side of the government's budget can be used to close a given sustainability gap.

196. This approach is also followed by the European Commission, which publishes for all member states the fiscal adjustment necessary to close the sustainability gap in its "**Sustainability Reports**". In these reports, the projections of future revenue and income streams take into account future outlays on pensions, healthcare, long-term care, unemployment benefits and education. Implicit government liabilities could of course also arise from policy changes in other areas, but the aforementioned expenditure items arguably have the most drastic impacts on the budget of future governments. While demographic trends vary between countries, there are some common trends in the development of projected expenditures. In general, projected expenditures on public pensions, healthcare, and long-term care increase, while expenditures on education and unemployment are assumed to decline. The same holds for potential growth rates, which are assumed to decline in the long run due to the ageing of societies. Table 14 documents how these public expenditure items are projected to change for Germany, France, and the aggregated EU 27.

Table 14

Age-related government-expenditure: 2010 and 2060[1)]

Percent of GDP

	Germany		France		EU 27	
	2010	2060	2010	2060	2010	2060
Pension spending	10.2	12.7	13.5	14.1	10.2	12.5
Healthcare	7.6	9.2	8.2	9.3	6.8	8.2
Long-term care	1.0	2.4	1.5	2.2	1.3	2.4
Unemployment benefits and education	4.6	4.2	5.8	5.6	4.9	4.7
Total	23.3	28.4	29.0	31.2	23.2	27.8
Change 2010 to 2060 (percentage points)	5.1		2.2		4.6	

1) Source: European Commission „Sustainability Report 2009".

Data

197. The reports present **two variants** of a sustainability indicator, named S1 and S2. Both indicators measure the permanent improvement of the primary structural balance – the structural balance excluding interest payments – needed to ensure fiscal sustainability. This definition of the fiscal balance allows a focused view on the underlying fiscal stance, independently of the economic cycle, temporary effects and predetermined interest payments. According to the first indicator, **S1**, fiscal sustainability is defined as the necessary adjustment of the primary structural balance required for reaching a **target debt ratio** of 60 per cent relative to GDP in the year 2060.

The second indicator, **S2**, is defined as the adjustment of the primary structural balance necessary to fulfil the intertemporal budget constraint over an **infinite horizon**. The necessary adjustment of the fiscal balance is calculated as percentage points of GDP. Thus, if the indicator S2 displays an adjustment need of, say, 3 percentage points, this means that public expenditures (revenues) have to be permanently reduced (increased) by 3 percentage points of GDP in order to reach a sustainable fiscal position. Alternatively, government could reduce its implicit liabilities, leaving the structural balance unchanged. If indicator S2 is positive, this implies that, if no corrective action is undertaken, the sum of the explicit and implicit public debt relative to GDP will explode in the long run, violating the intertemporal budget constraint. For the **dashboard** it seems appropriate to select the S2 indicator, not least because it is easier to calculate. Table 15 shows the main results for Germany, France and for the consolidated EU 27.

198. The S2 indicator can be calculated as the sum of **two components**. First, one has to estimate the adjustment that is needed to stabilize the debt-to-GDP ratio. And second, one needs to assess the additional adjustment requirements caused by rising expenditures due to an ageing population. The necessary adjustments are always expressed as the required percentage-point improvements in the primary structural balance.

For **Germany**, the indicator S2 documents an adjustment need of 4.2 percentage points of GDP in the primary structural balance (columns B + C in Table 15). This means – while holding the implicit government liabilities constant – that Germany would need to

improve its structural primary surplus from 0.6 in 2009 (as estimated in the EC's report at the time of publication) to 4.8 per cent in order to close the sustainability gap. Part of the adjustment could also result from reducing implicit government liabilities.

For **France**, the indicator displays an adjustment need of 5.6 percentage points. Starting with a structural primary deficit of -2.7 per cent of GDP in 2009, a primary surplus of 2.9 (= -2.7 + 5.6) per cent of GDP would be required to close the sustainability gap. These numbers show the importance of **demographic trends** for the sustainability of public finances. As the population in Germany is ageing faster than in France, the primary surplus required in Germany to close the sustainability gap (last column in Table 15) is higher than that in France. Nevertheless, the adjustment need, the change in the structural balance, is higher for France, due to its higher initial structural deficit.

Table 15

Fiscal sustainability calculations[1)]

Percent of GDP

	Structural primary balance 2009	Required improvement of structural primary balance due to…			Sustainable structural primary balance[3)]
		…stabilization of debt ratio	…additional age-related expenditure	S2 indicator[2)]	
	A	B	C	B + C	A + B + C
Germany	0.6	0.9	3.3	4.2	4.8
France	– 2.7	3.8	1.8	5.6	2.9
EU 27	– 2.0	3.3	3.2	6.5	4.5

1) Source: European Commission „Sustainability Report 2009".– 2) Necessary adjustment of structural primary balance required to close sustainability gap.– 3) Adjustment can also be achieved through cuts in implicit guarantees.

Data

199. These results show that while Germany's constitutional rule that sets an ambitious ceiling of 0.35 per cent of GDP for the central government's structural deficit is helping to contain explicit government liabilities, it is not sufficient to ensure the closure of the sustainability gap as calculated here after taking into account implicit government liabilities. **Institutional reforms** of the kind that led to anchoring budget rules in Germany's constitution are under review in other countries. In France, a commission chaired by Michel Camdessus presented propositions of this kind to the prime minister on 25 June 2010. The main propositions are to constitutionally enshrine, first, the necessity to decide fiscal and social expenditures solely through parliamentary financial regulations and, second, the obligation of a multi-year financial plan featuring a mandatory roadmap for reducing the deficits and attaining balanced public finances.

200. As with all calculations of this kind, a number of assumptions are necessary to derive these results. Estimates have to be calculated for, among others, life expectancy, labour productivity, potential output, real interest rates and future expenditures and revenues due to ageing. The EU presents **sensitivity tests** for some of these variables. Table 16 shows the adjustment need for different paths of potential output. In the **baseline scenario**, annual poten-

tial GDP growth is assumed to average 2.4 per cent for the years 2007–20. Thereafter, annual potential GDP growth is assumed to decline significantly on account of the shrinking working-age population, which acts as a drag on growth and on per capita income. By the 2041–60 period, GDP growth is projected to average 1.3 per cent per annum. The **permanent shock scenario** assumes that the growth rate of potential output never recovers from the crisis. In this scenario, the adjustment need is 1.6 percentage points of GDP higher for Germany and 2 percentage points of GDP higher for France compared to the baseline scenario.

Table 16

Calculations of S2 under alternative growth scenario[1)]

Percentage points of GDP

	Baseline scenario	Alternative growth scenario: permanent shock
Germany	4.2	5.8
France	5.6	7.6
EU 27	6.5	8.0

1) Source: European Commission „Sustainability Report 2009".

Data

201. To sum up, we suggest adding two indicators to the dashboard to monitor the sustainability of public finances.

- The **cyclically-adjusted public sector balance** (as reported by the European Commission) should not exceed net public investment (in line with the Golden Rule) or be even lower in the case of countries with stricter rules, especially given positive fiscal sustainability gaps.

- The **fiscal sustainability gap** (as represented by the S2 indicator in the European Commission's Sustainability Reports) should decrease over time and eventually converge towards zero in order to signal a sustainable fiscal stance. Two points should be noted in this context, however. First, to permit regular reporting on the dashboard, the European Commission would have to update the indicator every year. Second, it is important to bear in mind that this indicator is more sensitive to specific assumptions and projections than other indicators that have been selected for the dashboard.

3. Financial sustainability

202. Recent decades have repeatedly provided evidence that rapid credit growth and asset price booms might turn out to be unsustainable in the long run, with detrimental consequences for households, the corporate sector and financial intermediaries. In fact, excessive credit booms have regularly ended in **financial crises** and in large-scale destruction of wealth. Already in the boom phase or pre-crisis period, the misallocation of resources and investments can lead to welfare losses, as savings are channeled excessively into projects with a low or even negative rate of return. During this time, the measurement of GDP will typically be distorted upwards and will signal increases in welfare which are unsubstantiated in reality. The

crisis itself then acts as a corrective, redirecting the assessment of welfare towards its genuine level.

Unfortunately, the consequences of serious crises often extend beyond a mere correction since they tend to lead to a **permanent destruction** of human and physical capital. In addition, unsustainable private sector debt will frequently translate into a large increase in public sector debt since the public sector is compelled to put its own balance sheet on the line. Thus, in hindsight it would have been better if the regular measurement of economic performance and well-being had given a timely indication that an unsustainable situation was building up, allowing for corrective action before a full-fledged crisis could emerge. Yet, up to the present time, information on financial sustainability has not been part of the portfolio of economic indicators being regularly reported by statistical offices.

This section seeks to lay the ground for the regular documentation of the state of financial sustainability which should complement the monitoring of current economic performance and well-being. To this end, we discuss a **range of indicators** which signal unsustainable developments in the private and in the financial sector, which in many countries covers both private and public financial institutions. Notably, the intention of suggesting these indicators is not to address issues of business cycle smoothing. Instead, the objective is exclusively to investigate **excessive** fundamental and undesirable developments which are typically associated with boom phases that are likely to lead to severe economic crises such as the one we are currently experiencing.

203. While our objectives are ambitious, we have to remain realistic: it will never be possible to predict financial crises with certainty. What we might be able to offer, though, is a small set of reasonably robust **early-warning indicators** which could alert policy makers and the general public to fundamental undesirable developments in the financial sector. It goes without saying that this limited set of indicators should not be misconstrued as a substitute for detailed macro-prudential supervision, existing early-warning systems or any other methods already employed by experts and sovereign authorities – particularly prudential supervision – to analyze the sustainability of the economic environment. Moreover, it should be borne in mind that these indicators do not cover all relevant areas comprehensively. Rather, their purpose is the timely identification of economic developments which, if left uncorrected, might lead to distressed situations. If the indicators give **alert**, policy makers should consult experts and authorities, and, if alerts are confirmed, take preventive steps. The indicators might also serve as a control mechanism for the wider public, as they provide the basis for an informed discussion of financial sustainability.

204. This section is structured as follows. First we provide a brief argumentation for including indicators of private and financial sector financial sustainability in our dashboard and contrast our approach with the recommendations of the SSFC Report. Then we identify indicators for policy makers and the wider public.

Financial crises and sustainability

205. As a consequence of advancing **globalization** and integration, countries and markets alike are becoming more and more interlinked. This process has intensified particularly over recent decades. Especially European economies have experienced deep political and economic integration. This development creates substantial potential for economic growth, but it simultaneously increases the danger of international contagion (Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, 2009). Both recent and historical experiences as well as economic theory suggest that financial crises are costly and that they are hardly a rare event. Thus, it is advisable to incorporate early-warning signals into the regular reports by statistical offices.

206. The **crisis** of the past few years has shown just how costly financial crises can be. The IMF estimated global bank write-downs over the period 2007-2010 at US$ 2,810 billion worldwide (of which US$ 814 billion relates to euro-area banks) on bank holdings of both loans and securities – which amounts to a tremendous **destruction** of **financial assets** (IMF, 2009). The costs to the real economy can be gauged from the contraction of current production and employment of resources. In 2009, real **GDP growth** in the EU was negative with a rate of -4.2 per cent and the EU's **unemployment rate** increased to 8.9 per cent. Also, a recent study on systemic banking crises over the past four decades shows that cumulative **output losses** associated with banking crises can be substantial, averaging about 20 per cent of GDP during the first four years of the crisis (Laeven and Valencia, 2008). Additionally, crises in the private sector also affect public sector finances. Hence, cumulative **fiscal costs** of systemic banking crises can be large, too, averaging about 13.3 per cent of GDP, and can be as large as 55.1 per cent of GDP (Caprio and Klingebiel, 1996; Hoggarth et al., 2001).

207. Financial and economic crises are **not rare** (Bordo et al., 2001). Recent examples are the American Savings & Loan crisis in the 1980s, the northern European banking crisis in the late 1980s and early 1990s, the Asian crisis at the end of 1990s, and the dot-com crisis at the beginning of the 2000s. In fact, the particular causes and roots typically differ from crisis to crisis. Furthermore, their international propagation as well as their intensities vary (Caprio and Klingebiel, 1996). However, they have one thing in common: they are **highly disruptive** for the economies affected as economic performance slumps, unemployment surges and economies are thrown back in their development, sometimes by five to ten years.

Most importantly, as shown by Reinhart and Rogoff, large-scale financial crises are usually followed by a surge in **public debt** levels that often end in fiscal crises (Reinhart and Rogoff, 2009, 2010b). It clearly follows from this that avoiding the build-up of financial imbalances is crucial not only to prevent short-term volatility in macroeconomic aggregates. Preventing financial crises would automatically improve the measures of fiscal sustainability discussed above.

208. The frequency and severity of financial crises have been reflected in economic theory. According to Minsky, one of the most prominent contributions to this literature, financial fragility levels move in tandem with the business cycle (Minsky, 2008). In short, risks are built

up in boom phases and materialize in downturn phases. In an environment with persistent economic growth and rising expected profits, firms eventually start to engage in speculative financing. Although they are well aware that current profits will not cover all the interest charges, more and more firms believe that profits will rise steadily and the loans will eventually be repaid. More loans, in turn, lead to more investment, thereby further promoting economic growth. As lenders in turn are infected by the euphoria of high profits, they contribute to the cycle by lending even more. In this phase, many economic agents are ignoring or at least underestimating the accumulating risks. This spiral moves on up to a point at which the economy has taken on too much risky credit.

Then, it is only a matter of time before some large firm actually defaults. This is the point where the upturn comes to a halt and turns into a downturn. Lenders suddenly perceive the actual risks and become very conservative in giving credits. Refinancing becomes difficult or even impossible for many firms, with the consequence of more defaults. A real economic crisis with a downward spiral begins if no new credit resources are found to sustain the refinancing process.

As it is a common tendency to be overly euphoric regarding economic performance in boom phases, and risks are all too easily overlooked, early-warning indicators could be a suitable means for policy makers to reflect on themselves and for the wider public to reflect on policy makers. In fact, there were some highly renowned economists who, prior to the crisis, warned of the looming risks which were likely to materialize. However, their warnings were unheeded and they were ignored against the background of the outstanding economic performance worldwide.

209. The SSFC Report acknowledges the need to complement current measures of well-being and development with indicators that signal unsustainable developments in the private and financial sector. It points out that the current crisis has shown that "neither the private nor the public accounting systems were able to deliver an early warning". In particular, "some of the performance was a 'mirage', profits that were based on prices that had been inflated by a bubble." Against this backdrop, the SSFC argues that "metrics which incorporated assessments of sustainability (e.g. increasing indebtedness) would have provided a more cautious view of economic performance".

In terms of concrete action, the SSFC Report proposes the stress testing of balance sheets, with alternative valuations to take into account situations in which market prices for assets are not available or are subject to bubbles and bursts. However, as the SSFC Report was still in progress before the crisis was in full swing and had revealed the severity of these problems, it does not pursue a broader discussion on this issue. By including a set of indicators that aim to capture the problem more directly, we hope to be able to fill this gap.

Identifying appropriate indicators

210. In line with our dashboard approach, it is the goal of this study to provide policy makers with a limited set of **robust leading indicators** for financial distress. While these indicators

are intended to empower policy makers to draw conclusions regarding the probability of an impending crisis, they also have to be **parsimonious** enough to be a viable part of regular statistical reporting. Thus, these indicators cannot serve as a substitute for comprehensive and complex experts' tools. Consequently, the challenge here, as in the case of current material and non-material well-being, and of environmental and fiscal sustainability, is to reduce the wide range of available indicators to a limited set of robust entries which provide a valuable summary of financial developments.

The timely detection of imbalances which might lead to severe crisis if left uncorrected is hardly a trivial task. One needs to identify appropriate and robust leading indicators applicable in many circumstances, and yet crises vastly differ in their **concrete causes** and in the **initial conditions** framing the economic environment on their outbreak. The search for common characteristics, however, is not in vain as Kindleberger (1978) – in studying financial crises – notes: "For historians each event is unique. Economics, however, maintains that forces in society and nature behave in repetitive ways. History is particular; economics is general." Thus, there are grounds for hope as there is a growing literature dedicated to the identification of variables that have fairly robust properties as appropriate leading indicators of financial crisis prediction.

In economic research, the quest for indicators which are robust predictors of unsustainable developments is not new. In fact there is a **large literature** dealing with the optimal reaction of monetary policy to asset-price bubbles, the potential of leading indicators for predicting currency and financial crises, and non-parametric early-warning systems. There is also an emerging consensus on a few variables that have fairly robust indicator properties and which are increasingly being monitored by early-warning systems at central banks and in international organizations.

211. Broadly speaking, there is a bi-polar system of indicators. On the one hand, there are highly **aggregated indicators** which reflect an overall measure of various **disaggregated indicators**. While aggregated indicators are less granular, they are more manageable and can easily be understood by the wider public. On the other hand, there exists a wide variety of disaggregated indicators which could be scrutinized as well. Although these indicators are highly granular and, thus, allow deeper insights, they are too complex for our purposes and hence should be handled exclusively by experts and supervisory authorities. As we intend to provide policy makers and the wider public with a manageable and intuitive set of indicators, we prefer to choose from among aggregated indicators.

212. A vast empirical literature attempts to identify appropriate indicators with which the accumulation of risks and their materialization can be foreseen. A selection of this kind of literature is summarized in Table A1, pages 146 pp. Obviously, older literature has the tendency to use a wide range of various indicators for the investigation of financial distress, while recent literature tries to focus on a small set of indicators – which is also the purpose of this study – and on the objective of determining the most relevant indicators. In particular, some studies and advocate concentrating on a small and manageable set of variables (Borio

and Drehmann, 2009a; Borio and Lowe, 2002a, 2002b). Reviewing the empirical literature, **credit growth** and **asset prices** seem to crystallize as appropriate and commonly accepted indicators. Particularly against the backdrop of the current crisis, these two indicators appear to be a reasonable selection: credit growth and asset prices experienced an inflationary increase prior to the crisis – increasing far more than income.

One of the relatively few robust findings to emerge from the empirical literature on leading indicators of banking crises is that **rapid domestic credit growth** increases the likelihood of a problem (compare Table A1). This conclusion can already be drawn from early studies. A later result is that also rapid and **persistent increases** in **asset prices** contribute to the likelihood of financial distress. This is a robust finding which emerged during the recent decade when reliable data on asset prices became available. For instance, property prices covering a sufficiently long period were scarcely available before the BIS started to collect them in 1990 (Borio and Lowe, 2002a). Some empirical studies demonstrate that sustained rapid credit growth combined with large increases in asset prices appears to increase the probability of an episode of financial instability (Borio and Drehmann, 2009a; Borio and Lowe, 2002a, 2002b).

213. Hence, we recommend that policy makers and the wider public concentrate on the following three indicators:

- total private credit relative to GDP (both in nominal terms),
- real equity prices (deflated by the consumer price index),
- real property prices (deflated by the consumer price index).

While the dashboard should primarily consider these three indicators, a number of corresponding memoranda items should also be shown. In particular, total private credit to GDP should be broken down for the **non-financial** and the **financial** sector, while real property prices should be decomposed into **commercial** and **residential** property prices.

214. Generally, one must be careful in interpreting the level of credit growth as such. An increase in credit growth does not necessarily signal an overheating of demand. It may, for instance, be the result of improving supply-side conditions. Hence, **credit growth** should not be considered independently from income, i.e. GDP. If an increase in credit growth is accompanied by a similar increase in income, then we still might consider credit growth to be sustainable. On the other hand, if income increases as a consequence of the real effects of a credit bubble, there is cause for concern.

It follows that we must consider a threshold at which we would consider credit growth to be unsustainable. For instance, if credit grows in line with GDP, we obtain a more or less constant ratio of credit to GDP. If, however, credit growth is substantially and persistently higher than GDP, then we recognize a departure of the current ratio of credit to GDP from its long-run sustainable path. In practical terms, we identify an increasing **credit gap**, which might

signal an unsustainable level of indebtedness in the private sector (i.e. both non-financial and financial). The same method is applied to the asset price indices to derive both **equity** and **property price gaps** (see Box 4 for a more detailed discussion of methodological issues).

215. It would also be appropriate to use ratios – where the denominator refers to the real economy – for analyzing **asset price** indicators. This would be consistent with the general idea that growth in credit and asset prices should not be considered in isolation from some measure of income to finance these developments. More specifically, one could apply the equity prices-to-earnings ratio and property prices-to-rental ratio. However, two constraints have to be considered. First, the availability of data has to be ensured – not only for France and Germany on which we focus here, but also for a number of other countries – so that an extensive cross-country analysis is possible. Second, it is necessary to test how sensitive asset price ratios are when evaluating their early-warning properties (see Box 4). Hence, for the time being, we suggest using the above-mentioned three indicators (one as a ratio and two as indices) supported by the research from Borio and Lowe as well as Borio and Drehmann, who show that these early-warning indicators generate robust results (Borio and Drehmann, 2009a; Borio and Lowe, 2002a, 2002b).

Box 4

Methodological issues

Following Borio and Lowe and Borio and Drehmann, we focus on three core variables which – a claim also supported by the literature in the appendix – supposedly contain useful information about the development of financial imbalances: the ratio of (private sector) credit to GDP, real equity prices, and real (commercial and residential) property prices (Borio and Drehmann, 2009a; Borio and Lowe, 2002a, 2002b).

As both considerable credit growth and rising asset prices are hardly unequivocal warning signs, we need to translate their development into a set of leading indicators in a way that reflects crisis potential. To this end, we follow the so-called gap approach. This approach tries to capture the cumulative processes that in the boom phase sow the seeds of subsequent distress by employing deviations of the core variables (measured in levels) from an estimated trend. Typically, the trend estimate might be implemented by a Hodrick-Prescott filter, but detrending methods other than the HP filter may be used instead, e.g. linear filters. In the current context, however, the HP filter – notwithstanding its statistical flaws – has proved to generate robust results.

Subsequently, the deviations in each period are added up to so-called “gaps” to account for the cumulative processes which occur over the short and medium run, i.e. from about one to five years. For instance, the HP filter is applied to real property prices, producing an estimated trend or “filtered” series. Then, for each period the deviations of actual real property prices from the filtered series are recorded and summed. Following this approach, “gaps” are building up whenever moderate and persistent above-trend developments are observed and also when above-trend developments surge spontaneously.

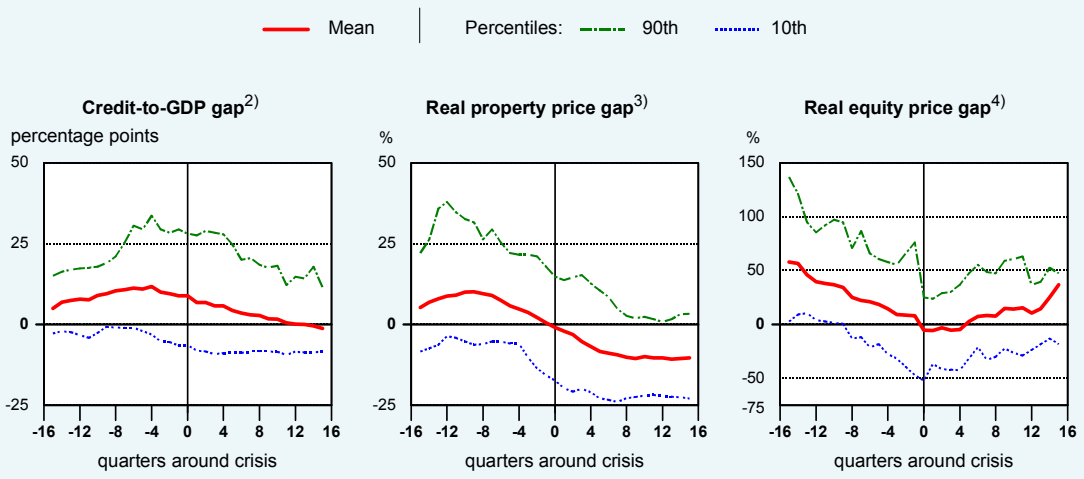
Now it is one thing to assess the build-up of gaps in hindsight, when the detrended series incorporates information on the further development of the series. Yet, authorities have to make their assessments in “real time”. That is, the gaps which they will be able to calculate in practice can only incorporate the information available at the time the assessments are made. Consequently,

they have to be based on one-sided trends. Thus, as is the case in all other instances of trend estimation which support the regular monitoring of economic performance, in practical work the trend has to be re-estimated, and the gap calculation revised, whenever more recent data become available.

Generally, asset price misalignments are captured by asset price gaps, while systemic shock absorption capacity is approximated by credit gaps – representing a rough measure of leverage for the economy as a whole. The expectation is that if the credit-to-GDP ratio, real equity prices or real property prices move "sufficiently above" their trend, i.e. they exceed some critical threshold, then financial imbalances are emerging, signaling the risk of subsequent financial distress (Borio and Drehmann, 2009a).

Chart 26

Credit and asset price behaviour around banking crises[1)]



1) The historical dispersion of the relevant variable is taken at the specific quarter across all crisis countries. Gaps are estimated using a one-sided rolling Hodrick-Prescott filter with lambda set to 400,000. The gaps are calculated for a sample covering 18 industrialized countries (Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, the United Kingdom and the United States) over a period from 1980-2003.– 2) In percentage points as deviations from trend.– 3) Weighted average of real residential and commercial property prices with weights corresponding to estimates of their share in overall property wealth; the gap is in per cent relative to trend.– 4) Equity prices are measured in real terms; the gap is in per cent relative to trend.

Data

Source: Borio and Drehmann (2009a)

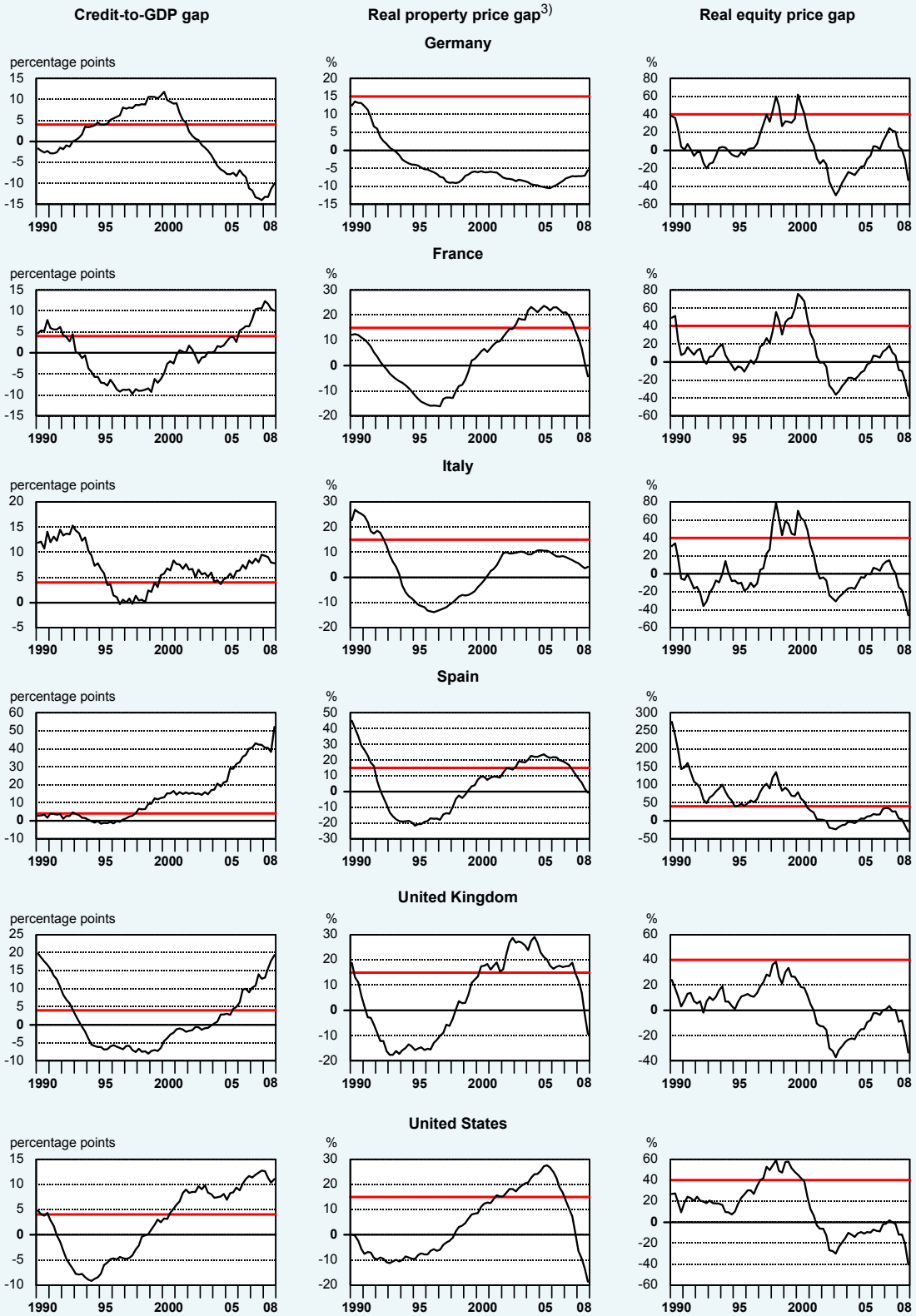
Chart 26 shows the average credit gap, equity price gap and (commercial and residential) property price gap around crises events. For the crises events, the standard definition of banking crises employed in previous research is taken (Borio and Drehmann, 2009a). It shows that – on average – credit, property and equity price gaps tend to be large and positive prior to crises. In addition, the property and equity price gaps peak well before the crisis, with those of equity prices peaking before property prices and being much larger. By contrast, the credit gap peaks only one year in advance of the crisis. At the same time, all three indicators exhibit considerable dispersion.

To use these variables as leading indicators for crisis episodes, a threshold value has to be defined; a crisis is then likely to occur if the indicator exceeds the threshold. Applying this method yields estimates of the optimal thresholds at 4 percentage points for the credit-to-GDP gap, 15 per cent for the property price gap, and 40 per cent for the equity price gap (Borio and Drehmann, 2009a). The underlying condition for the definition of the optimal thresholds is to minimize the noise-to-signal ratio subject to predicting at least three quarters of the crises.

Chart 27

Estimated cumulated gaps[1)]

Threshold[2)]



1) Calibration is in-sample, from 1980 to 2003, while the out-of-sample exercise is performed for the years 2004 to 2008.– 2) The threshold is 4 percentage points for credit-to-GDP gap; 15% for real property price gap and 40% for real equity price gap.– 3) Weighted average of residential and commercial property prices with weights corresponding to estimates of their share in overall property wealth. The legend refers to the residential property price component.

Source: Borio and Drehmann (2009a)

Data

The noise-to-signal ratio is defined as the ratio of the frequency of Type I errors – i.e. the percentage of non-crisis periods in which a crisis is incorrectly signaled – to one minus the frequency of Type II errors – i.e. one minus the percentage of crises that are not correctly predicted. According to Borio and Drehmann, this objective should provide a good balance between identifying costly crises and missing them (Borio and Drehmann, 2009a). Note that the thresholds and underlying optimization function can vary according to the user's objective function – e.g. with lower thresholds capturing a larger percentage of crises, but at the expense of a higher noise-to-signal ratio.

Given these threshold values, the general performance of the indicators is quite good. At a three year horizon, about three quarters (77 per cent) of the crises are predicted with a noise-to-signal ratio of less than 20 per cent. This means that, for every ten signals issued, about two incorrectly point to a crisis. Note that these encouraging results refer to in-sample calculations. At the same time, although out-of-sample predictions are less exact, they are still promising as more than 50 per cent of the crises can be predicted with a noise-to-signal ratio of less than 70 per cent. (Although the noise-to-signal ratio increases considerably compared to in-sample estimates, this can partly be explained by the low number of "non-crisis" periods over the period 2004 to 2008, which can lead to large movements of the noise-to-signal ratio in response to small changes in the absolute number of Type II errors.)

To obtain a more concrete impression of the performance of the leading indicators, some case studies can be considered (Chart 27). In fact, it is straightforward to contemplate the behavior of the leading indicators with respect to the current crises. The data show that the credit-to-GDP gap would have indicated an excessive development prior to the current crisis in most of the cases, with Germany as an exception. Similarly, the property price gap would have given alarm in many cases. By contrast, the equity price gap would have completely failed to signal the build-up of risks – regarding the current crisis. Note that the gaps would have indicated vulnerabilities already at the beginning of the 1990s and also around 2000 – with all three indicators delivering relevant information.

Generally, indicators can be used separately for the prediction of financial distress. Note that combinations of indicators do not necessarily increase the number of correctly predicted crises, but typically reduce the number of wrong signals (noise) and hence the noise-to-signal ratio. This particularly applies to combinations of credit growth and asset prices, i.e. credit-to-GDP ratio and either property prices or equity prices.

216. This gap approach fits in with the general idea that unsustainable developments have a cumulative effect, i.e. vulnerabilities generally build up over an extended period, rather than in a single year. A large gap could develop through either one year of very rapid credit growth, or alternatively as the result of a number of years of above-trend growth. Similarly, asset price booms are defined as periods in which real asset prices deviate from their trend by specified amounts, indicating an **asset-price gap** (Borio and Lowe, 2002a). The crucial question, however, is when is the gap large enough to be considered unsustainable? In other words, what is the crucial threshold?

According to the literature, a common procedure is to determine the **threshold** so as to minimize the so-called noise-to-signal ratio, which is the number of wrong signals relative to the number of correct signals given by the indicator (Kaminsky and Reinhart, 1999). For instance,

a noise-to-signal ratio of 1.0 indicates that each correct signal is accompanied by a false signal (see Box 4). Following Borio and Drehmann, we suggest applying a threshold of **4 percentage points** for the **credit growth** indicator, **15 per cent with respect to** the **property price** indicator and **40 per cent** regarding the **equity price** indicator (Borio and Drehmann, 2009a).

217. As Borio and Drehmann point out, minimizing the noise-to-signal ratio is just one of several options for specifying the **objective function** that guides the identification of concrete thresholds (Borio and Drehmann, 2009a). In particular, one could place more weight on detecting impending crises successfully – even if this implies increased noise, i.e. an increase in false signals. In this case, lower thresholds apply, say, 3 percentage points for credit growth, 10 per cent for property prices and 30 per cent for equity prices. It must be clear, however, that the role of this trade-off crucially depends on the objectives of the recipient of the information. For instance, the wider public might prefer a high detection rate of actual crises, accepting a relatively high noise, since it might not be very costly for the public to receive a relatively large number of wrong signals. By contrast, this might be very costly for authorities, as they would take appropriate action in anticipation of an impending crisis. If the signal were wrong, efforts and economic resources would be wasted.

4. Environmental sustainability

218. Environmental sustainability is the third, and arguably the most prominent, of the three essential aspects distilled in the academic and public discussion about what constitutes sustainable development. In particular, an environmentally sustainable system must not squander its resource base. This can only be achieved if one **avoids** an **over-exploitation** of **renewable resources** or of environmental sink functions. One should ascertain, for instance, the absorption capacity for carbon dioxide by oceans or forests, and prevent groundwater depletion due to the over-consumption of water. Environmental sustainability also requires that the **depletion** of **non-renewable resources** is managed in an efficient and intergenerationally equitable way. Finally, environmentally sustainable systems require the **maintenance** of **biodiversity** to ascertain the system's resilience to shocks (Polasky et al., 2005).

In this section, we review and, keeping an eye on the trade-offs involved, evaluate existing indicators that cover these problems as **candidates** for our dashboard. We decided, first, to include two indicators of **greenhouse gas emissions** in our comprehensive dashboard, one expressed in terms of levels and one reporting emissions per capita. As the water problem could not be meaningfully treated from a national perspective, however, we do not pursue this issue further. Second, we also suggest two indicators of **resource productivity** and consumption for inclusion in our dashboard. Finally, despite some reservations regarding its appropriateness, we also decided to include one preliminary **biodiversity indicator**. We are aware that the ideal setup of these resource indicators, and especially that of the indicator of biodiversity, cannot be determined by economists alone. Our choices therefore are intended to reflect the current, rather unsatisfactory state of the discussion as well as possible. We are completely open, however, to adjusting the dashboard entries once ongoing interdisciplinary research provides better guidance to more appropriate measures.

The need to monitor environmental sustainability

219. The broad definition of environmental sustainability which we follow in our work accords with the arguably worldwide consensus that the most pressing **environmental problems** are climate change, the depletion of non-renewable resources, the over-exploitation of renewable resources, and the ongoing loss of biodiversity. The extent to which these topics have taken centre stage in the political discussion is reflected by the **international agreements** passed during the United Nations Conference on Environment and Development in Rio de Janeiro in 1992 (Rio Summit). These comprise the Rio Declaration on Environment and Development, which consists of 27 principles intended to guide future sustainable development around the world, the Framework Convention on Climate Change (FCCC) and the Convention on Biological Diversity (CBD).

Whereas the Rio Declaration is only a short, non-binding document, both the FCCC and the CBD are legally binding agreements. The objective of the FCCC is **to stabilize greenhouse gas** concentration in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. This Framework resulted in the **Kyoto Protocol**, which commits the industrialized countries to reduce their greenhouse gas emissions. The CBD, finally, has three main goals: (i) conservation of biological diversity (**biodiversity**), (ii) sustainable use of its components and (iii) fair and equitable sharing of benefits arising from genetic resources.

220. In the aftermath of the Rio Summit, coordinated European and national strategies have been developed to implement these agreements. A central prerequisite for ascertaining that these various strategies are successfully contributing to achieving their ambitious targets is a **reliable monitoring** of the state of affairs, carried out at **regular intervals**. To this end, many research institutions and government agencies have drawn up a multitude of individual indicators as well as composite indicators, and various other aggregated measures. Clearly, environmental sustainability touches upon the key issues of economics, **scarcity** and the competition of wants. It follows from this that we will concentrate on those indicators out of this variegated reservoir that tend to emphasize these aspects, not least because we feel relatively competent in assessing the quality of such indicators.

Most of this work has been conducted **outside** of the **realm of economics**, for the obvious reason that developing the best indicators for the most important aspects of environmental sustainability also requires competence in the natural sciences as well as in social sciences other than economics. Consequently, in our own work we will necessarily have to rely on indicators whose selection or construction rest on the critical scientific discourse in other disciplines. Our specific choices presented here will therefore be contingent on the humble insight that they will be up for critical discussion by researchers from other disciplines and perhaps might even experience revision in the aftermath of this report.

Nevertheless, in our view economics has an **important contribution** to offer to the debate on environmental sustainability as well, epitomized in the insight that **trading off** the **welfare** of different individuals, let alone of different generations, can never be anything else but a

highly contentious issue. Thus, one of the iron-clad principles that economics adheres to in any discussion on welfare comparisons is that it needs to be clear whose welfare gains and losses are traded off against one another. Only then will one be able to discuss what requirements need to be fulfilled precisely in order to allow for such a trade-off. If this **minimum requirement** cannot be ascertained, there is no meaningful way of interpreting presumable indicators of scarcity – and one should rather not present them at all.

221. From the vantage point of economics, one would hope that it were possible to construct an overall indicator of the sustainability of a country's growth path by simply adjusting net investment in the **national accounts** to reflect the current treatment of natural resources. Indeed, while the standard model of optimal growth emphasizes the accumulation of physical capital as an engine of economic growth (Dasgupta and Heal, 1974; R. M. Solow, 1974), it can easily be augmented by additional stocks of production factors, such as non-renewable and renewable natural resources or human capital. Given an appropriate specification of society's intertemporal objective function (d'Autume and Schubert, 2008), one can derive the amount of **"genuine" savings** (or "adjusted net savings") as the sum of the net investments in physical, human and natural resources capital (Pearce et al., 1996; Hamilton and Clemens, 1999). Genuine savings can be interpreted as an indicator of sustainability, since a negative value indicates that the growth rate of well-being will inevitably become negative in the future (Hartwick, 1977).

Following this reasoning, the World Bank has computed adjusted net savings for 140 countries since 1990. To arrive at this indicator, the difference between gross national savings and fixed capital consumption is **augmented** by a measure of education expenditures and **diminished** by imputed values of natural resource depletion, damages caused by greenhouse gases and by pollution particles. This procedure looks appealing at first glance, but it faces severe difficulties in practice. Most importantly, the **prices** necessary to construct these imputed values are not at all easy to come by. While some approaches exist to derive appropriate values, such as contingent valuation, their implementation poses considerable problems. This is the principal reason why the SSFC report is **very critical** in its discussion of the concept of net adjusted savings.

Moreover, we have already emphasized net investment as an indicator for our dashboard in our discussion of growth sustainability. This measure is highly correlated with adjusted net savings, irrespective of the concrete adjustment procedure chosen. Thus, using genuine savings we could at best provide a slight variation of the information already included.

222. Consequently, we have therefore – once again – to decompose the task into fragments which are addressed one at a time. The first issue which we discuss in detail in this section is the problem of **climate change**, which is undoubtedly the best explored dimension of environmental sustainability both inside and outside of our discipline. Even for this well-researched problem, from the vantage point of economics the appropriate policy conclusions remain **disputed**, since there is simply no policy option that solely yields benefits for all future generations without imposing costs on the current generation. Moreover, global warming

tends to affect every region of the world, and climate effects of greenhouse gases do not depend on where the emissions occur. Even so, we feel fairly **comfortable** about interpreting the development of national greenhouse gas emissions since we know what accompanying information we need regarding global emission developments and concerning the options emerging for future generations when we assess national developments.

By contrast, research on **resource productivity** and, a fortiori, on the economics of ecosystems and **biodiversity** is still **work in progress** and therefore has not yet identified sufficiently the trade-offs on which any political discussion needs to focus. Furthermore, we know that both resource productivity and biodiversity will also have to be discussed on a global scale. In particular, where species are endangered, it is arguably of minor importance whether the endangered species live, say, in Germany or in France. Notwithstanding the fact that a subset of biodiversity services is provided by local ecosystems, it is not perfectly clear how national resource productivity or biodiversity indicators have to be embedded into this global discussion. Thus, given this imperfect state of the art, the selection of indicators we propose for monitoring resource productivity and the loss of biodiversity is clearly meant to be **preliminary** and open to future revision.

Greenhouse gas emissions

223. According to the current state of knowledge, rising levels of carbon dioxide and other greenhouse gases in the atmosphere have already caused **global warming** and will induce climate change on an even broader scale. The **consequences** of this global warming include rising sea levels, an increase in extreme weather events, acidification of the oceans as well as an accelerated loss of species and ecosystems. Furthermore, climate change might jeopardize water supply and food production, pose additional health risks, intensify conflicts and accelerate migration. Thus, climate change has the potential to trigger major social and economic crises. It is certainly difficult to condense these multifarious negative consequences into a single quantitative figure. The Stern report provides such an attempt and estimates that **damage** due to extreme weather events resulting from a global rise in temperature of about 2°C lies in the range of 0.5 % to 1 % of world GDP per annum. However, the problem might turn out to be even more severe since, according to the "business as usual" scenario of the Stern report, global temperatures could even rise by more than 5°C in the decades after 2100, at a correspondingly much higher economic cost (Stern, 2007).

Assessments such as these have encouraged international agreements to curb climate gas emissions. Most recently, in the **Copenhagen Accord** adopted at the UN Climate Change Conference in December 2009, the vast majority of countries agreed that the increase in the mean of global temperatures should be kept below 2°C. Environmental specialists reckon that, to realize this objective, cumulative CO_2 emissions until 2050 would have to be capped at 750 Gt (Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen, 2009a).

224. Limiting greenhouse gas emissions requires international climate agreements such as the **Kyoto Protocol**, which stipulates **concrete emissions targets** for participating nations and, thus, a cumulative objective. Currently, a total of 190 countries have signed this Proto-

col, agreeing that greenhouse gas emissions should be reduced by 5.2 % between 2008 and 2012 compared with emissions in the reference year 1990. The EU 15 countries committed themselves to reducing greenhouse gas emissions by 8 % on average. While the objective of **France**, given its low per capita emissions, is the stabilization of its emissions, with a reduction target of 21 %, **Germany** accepted the highest reduction volume of any individual country. Other heavy polluters like China, which ratified the Kyoto Protocol in 2002, are exempt from any reduction commitments, however. And still other heavy polluters, like the United States, have not yet even signed the Protocol.

Quite frustratingly for the proponents of climate protection, in December 2009, the UN Climate Change Conference in Copenhagen failed to decide on a comprehensive and binding follow-up climate treaty. Nevertheless, as any such follow-up agreement – irrespective of whether and when it comes about – would necessarily involve the specification of **national emission levels** of the participating countries, it seems **sensible** to include an indicator of greenhouse gas emissions in our dashboard. After all, Germany, France and the EU at large will hardly tend to compromise their roles as "climate pioneers" which persist in pursuing ambitious carbon dioxide reduction targets even if other countries do not join the club. After the summit in Copenhagen, the European Union maintains committed to reducing at least 20 % of its emissions by 2020 compared with the reference year 1990. But the French, German and British governments would like to convince their European partners to increase this reduction target to 30 %.

225. We can base our endeavour to monitor national emission levels on **comprehensive efforts** undertaken by environmental specialists and government agencies. Both the French and the German National Strategy for Sustainable Development and – at the European level – the European Union Sustainable Development Strategy (EU SDS) report trends in man-made emissions of the **six greenhouse gases** (GHG) regulated by the Kyoto Protocol (carbon dioxide, methane, nitrous oxide, and the so-called F-gases hydrofluorocarbons, perfluorocarbons and sulphur hexafluoride). In these figures each gas is weighted by its global warming potential and aggregated to document total greenhouse gas emissions in CO_2 equivalents (Eurostat, 2007). To facilitate easier international comparisons, including Non-Annex I-countries such as most emerging economies, the OECD, on the other hand, reports emissions of **carbon dioxide** in its Factbook 2010, because CO_2 makes up by far the largest share of greenhouse gases.

Of course, the most relevant figure for climate change is the GHG total, expressed in level terms. Countries that are parties to the FCCC regularly submit national GHG inventories to the FCCC secretariat. These data are currently available from 1990 to 2008. To address the most relevant figures directly, we propose to report **total emissions of GHG** in the sustainability part of our dashboard, using FCCC data.

226. Correspondingly, Table 17 documents levels of GHG emissions for France and Germany for the years 1990, 2000 and 2008. Germany's GHG emissions were 958 million tons in 2008 and 1,232 million tons in 1990. The emissions of France were 27 million tons in 2008

and 563 million tons in 1990. While the emission levels are the relevant figures affecting climate change, for policy purposes it is necessary to relate these levels to the **national targets**. These are typically formulated as emission reductions, expressed as percentages of 1990 values. In 2008, Germany had reduced its GHG emissions by more than 22 % compared with 1990, which already satisfies the reduction target of 21 % between 2008 and 2012. For France the decrease is approximately 6 %, which similarly achieves its target. Thus, if the success of national climate policy were to be gauged in terms of national emission levels alone, these figures would seem quite satisfactory.

Table 17

Emissions in Germany and France

	Germany			France		
	1990	2000	2008	1990	2000	2008
	Million tons					
Greenhouse gas emissions[1)]	1,232	1,025	958	563	557	527
CO_2 emissions from burning oil, coal and gas for energy use	950	827	804	352	377	368
	Tons per capita					
Greenhouse gas emissions[1)]	15.5	12.5	11.7	9.7	9.2	8.2
CO_2 emissions from burning oil, coal and gas for energy use	12.0	10.1	9.8	6.1	6.2	5.7

1) The annual greenhouse gas (GHG) emissions are estimated and reported under the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol and the Decision 280/2004/EC. The so called Kyoto basket includes carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), perfluorocarbons (PFCs), hydrofluorocarbons (HFCs) and sulphur hexafluoride (SF_6). The different greenhouse gases are weighted by their global warming potential, and the results are expressed in CO_2 equivalents.

Data

Sources: IEA, OECD, UN

227. Environmental policy, however, is necessarily a global affair. We would not do justice to the issue if we were simply to report the national figures or if we unequivocally hailed adherence to national targets as a success. By contrast, national emission levels have to be supplemented by and embedded into a **global context**. Thus, we have to use the CO_2 emissions from burning oil, coal and gas for energy use reported by the OECD (Table 18, page 134). In 2008, the share of these emissions in GHG reported by FCCC was more than 80 % in Germany and approximately 70 % in France. In the base year 1990 Germany's share of CO_2 emissions compared with worldwide emissions was only 2.7 % and that of France 1.3 %. Correspondingly, the reductions that occurred between 1990 and 2008 only account for a negligible share of total worldwide CO_2 emissions. But far more importantly, other countries and the world as a whole **increased** their CO_2 emissions between 1990 and 2008 to a considerable extent. It remains an open question whether this is despite or even due to – as a market reaction ("carbon leakage") – European efforts, but the conclusion in our context is unequivocal. Because climate change is a global phenomenon, a national indicator of GHG emissions, considered in isolation, could be highly misleading. Thus, in our dashboard it should always be **complemented** by some summary figures documenting total GHG emissions.

Table 18 (page 134) documents the development of CO_2 emissions from burning oil, coal and gas for energy use worldwide and for different regions and countries. In 2008 29,381 million tons of CO_2 were emitted worldwide. This is an increase of 40 % relative to 1990. The EU 27 reduced their CO_2 emissions by 5 %; the OECD countries emitted 12,630 tons in 2008, which is an increase of approximately 14 %. The emissions of China increased threefold. The United State increased its emissions by almost 15 % and India emitted 591 million tons of CO_2 in 1990 and 1,428 million tons in 2008.

228. Obviously, an appropriate strategy limiting global anthropogenic GHG emissions requires a binding international agreement. Key elements of such an agreement should be a legally **binding target** of greenhouse gas emissions, an international emission **trading system** and an **allocation mechanism** that distributes emission allowances among the participating countries (Tirole, 2009). Predictions regarding the appropriate cap on global GHG emissions would ideally be based on the IPCC proposal, which was recently confirmed in the Copenhagen Summit, that global warming should be kept below 2°C compared with the pre-industrial level. Recent estimates suggest that, in order to achieve this target with a probability of two thirds, the global CO_2 budget should not exceed 750 Gt for the period from 2010 to 2050 (Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen, 2009b, 2009c).

Notwithstanding the considerable uncertainty associated with such estimates, once the global budget has been determined, it needs to be distributed among all countries. Even though quite different allocation mechanisms are conceivable, the **principle of equality** seems to be a good starting point for a fair distribution of the global budget. Thus, equal **per capita emission rights** all over the world would arguably form a sensible basis for the allocation of national emissions budgets – perhaps modified by some rules considering the high GHG emissions per capita which the developed countries have recorded in the past.

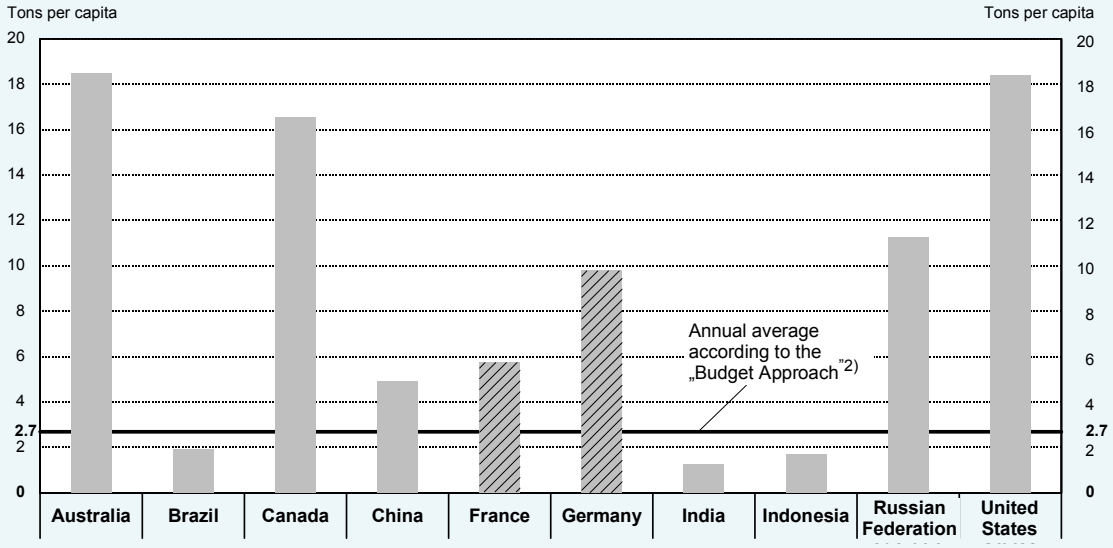
Based on a global budget of 750 Gt CO_2 until 2050 and on a projected world population of 6.9 billion in the year 2010, the CO_2 budget per capita would be 109 tons for the period from 2010 to 2050 or 2.7 tons annually until 2050. Compared with the current CO_2 emissions per capita in different countries, and abstracting from the possibility of achieving national emission targets by outsourcing the most polluting industries, it is obvious that achieving this target would require **tremendous reduction efforts** by the developed and newly industrializing countries (Chart 28). If a global emission trading system is established, however, these countries could buy emission allowances from developing countries.

229. Irrespective of its potential role in an allocation mechanism for globally traded emission permits, it would make sense to inform policy makers and the general public about national **GHG emissions per capita**. Thus, we propose including the current GHG emissions per capita as a second GHG indicator in our dashboard. Table 17 documents per capita emission levels of GHG for the years 1990, 2000 and 2008. In Germany, GHG emissions per capita were 11.7 tons in 2008, which is a reduction of nearly 25 % compared with 1990. In France, the

decrease in GHG emissions per capita over the same period amounted to approximately 15 % and the emissions per capita were 8.2 tons in 2008.

Chart 28

CO_2 emissions from burning oil, coal and gas for energy use in selected countries: 2008[1)]



1) Reported by the IEA (International Energy Agency) respectively by the OECD.– 2) Proposed by the German Advisory Council on Global Change (Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (2009b). It is based on the assumption that the global CO_2 budget should not exceed 750 Gt for the period from 2010 to 2050 if global warming should be kept below 2° C compared to the preindustrial level. The principle of equality is seen as a good starting point for a fair distribution of the global CO_2 budget. Thus, the annual average CO_2 emissions per capita could be calculated.

Data

230. Any meaningful discussion of these facts needs to address an issue which extends beyond the difficulties of achieving a global and binding climate treaty, namely the enforceability of any rule allocating pollution rights to different countries, or the power that some key actors, such as the United States or the BRIC states, wield in blocking any such climate treaty. As the costs of environmental damage are quite uncertain and their estimates disputed, and since policy inertia tends to be severe, there is an incentive, as with other collective goods, for any single country to **free-ride**. Hence, a crucial challenge for policy makers is to decide what to do if global and binding climate conventions are out of reach, or isolated emission reduction commitments are ineffective (Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2010). In such a scenario **adjustment strategies** to climate change may be a more viable alternative to active emission reduction policies. These arguments will have to be discussed in the light of current theoretical insights and empirical evidence whenever the dashboard is presented.

Resource productivity and resource consumption

231. One question that has been a focus of the political discourse on sustainability has been whether current modes of production lead to an **over-use** of present stocks of **natural resources**. Consequently, the intensity of resource extraction has so far played a prominent role in sustainability reporting. In particular, both the French, German and European sustainability strategies assume that an increase in resource productivity (of non-renewable resources) is a

policy goal worth pursuing. The German National Strategy for Sustainable Development, for example, requires that resource productivity in Germany should increase by 100 % between 1994 and 2010.

Conceptually, there are substantial differences between the monitoring of **renewable** and **non-renewable** resource use, which precludes a summary treatment under one and the same heading. For renewable resources, the main question is whether current extraction rates and environmental pressures endanger continuous reproduction. Answering this question usually requires close inspection of a given resource, for example fishing stocks or drinking water. Elaborate processes and regulations have to be set up to ensure constant monitoring and the regulation of respective industries. These have to distinguish between different fields, such as water, forests, fishing grounds and other harvesting products (see, for example, EU Commission, 2007). Without an **elaborate** societal and interdisciplinary **discourse** we cannot specify an order of priority in which to address the issue of renewable resources. Overall, the monitoring of specific renewable resources is a highly important endeavour, but it constitutes a research agenda on its own merits. Moreover, our discussion of biodiversity in the next subsection tends to overlap to some extent with the sustainability of renewable resources.

232. For these reasons, the following discussion focuses on **non-renewable resources**. Existing sustainability strategies use "resource productivity" as the key measure of sustainable resource extraction, relating total output to the total amount of a particular input. In the context considered here, it captures the amount of real GDP that can be **generated per unit** of (non-renewable) resources. This reflects the idea that a narrow focus on traditional measures, such as extraction rates and remaining proven reserves, cannot adequately capture the complexity of the issue.

The goal of fostering sustainability can be approached in two different, non-exclusive ways. First, raising the **recycling** ratio will reduce the speed of extraction of the stock of natural resources by enhancing the range of the stock which has already been extracted. The actual capacity of recycling for achieving this objective is quite uncertain, however. Second, enhancing production technology might increase overall resource **productivity**, which in turn will increase the range of existing stocks. One should not forget, however, that the effects of efficiency-enhancing technology improvements are typically offset to some extent by a demand for more output, due to the well-documented **rebound** effect. The principal mechanism behind this insight concerns the role of relative prices of the services generated by using the resource. Efficiency improvements save on resources for a given output, but simultaneously tend to make this output less costly to acquire, thereby increasing the amount of output demanded by consumers. Thus, it would be desirable to find an indicator of scarcities in the making which accounts for these intricacies.

Table 18

CO_2 emissions from burning oil, coal and gas for energy use in the world and by countries[1)]

Million tons

	1990	2000	2002	2003	2004	2005	2006	2007	2008
World	**20,965**	**23,497**	**24,070**	**25,111**	**26,357**	**27,129**	**28,024**	**28,945**	**29,381**
Australia	260	339	359	361	372	389	394	387	398
Austria	56	62	68	73	74	75	72	69	69
Belgium	108	119	112	120	117	113	110	106	111
Brazil	194	302	309	302	320	326	331	345	365
Canada	432	533	533	556	554	559	544	571	551
Chile	32	54	53	55	62	63	65	72	73
China	2,211	3,038	3,309	3,830	4,548	5,068	5,608	6,032	6,508
Czech Republic	155	122	117	121	122	120	121	122	117
Denmark	50	51	52	57	51	48	56	51	48
Estonia	36	15	14	16	17	17	16	19	18
Finland	54	54	62	72	67	55	67	64	57
France	**352**	**377**	**376**	**385**	**385**	**388**	**380**	**373**	**368**
Germany	**950**	**827**	**833**	**842**	**843**	**811**	**823**	**801**	**804**
Greece	70	87	90	94	93	95	94	98	93
Hungary	67	54	55	57	56	56	56	54	53
Iceland	2	2	2	2	2	2	2	2	2
India	591	981	1,021	1,046	1,117	1,160	1,250	1,338	1,428
Indonesia	141	268	293	299	314	324	339	365	385
Ireland	30	41	42	41	42	43	45	44	44
Israel	33	55	59	61	60	60	62	65	63
Italy	397	426	435	452	453	457	458	441	430
Japan	1,064	1,184	1,205	1,213	1,212	1,221	1,205	1,242	1,151
Korea	229	421	445	448	469	468	477	490	501
Luxembourg	10	8	9	10	11	11	11	11	10
Mexico	265	346	353	361	368	390	397	418	408
Netherlands	156	172	178	183	185	183	178	177	178
New Zealand	22	30	32	33	33	33	34	32	33
Norway	28	34	34	37	38	36	37	38	38
Poland	344	291	280	291	295	293	305	304	299
Portugal	39	59	63	58	60	63	56	55	52
Slovak Republic	57	37	38	38	37	38	37	37	36
Slovenia	13	14	15	15	15	16	16	16	17
Spain	206	284	302	310	327	340	332	344	318
South Africa	255	299	295	321	338	331	332	343	337
Sweden	53	53	54	55	54	50	48	46	46
Switzerland	41	42	41	43	44	44	44	42	44
Russian Federation	2,179	1,506	1,494	1,531	1,513	1,516	1,580	1,579	1,594
Turkey	127	201	192	202	207	216	240	265	264
United Kingdom	549	524	522	534	534	532	533	521	511
United States	4,869	5,698	5,605	5,680	5,758	5,772	5,685	5,763	5,596
EU 27 total	4,054	3,831	3,877	3,994	4,005	3,973	3,988	3,930	3,850
OECD total	11,045	12,476	12,490	12,730	12,863	12,903	12,841	12,970	12,630

1) Reported by the IEA (International Energy Agency) respectively by the OECD.

Data

233. The starting point in our quest for an indicator of the sustainability of non-renewable resources is a fundamental insight of economics. Barring **market failures**, the allocation resulting from the interplay of supply and demand leads to **efficient outcomes**. In the case of

non-renewable resources, an intertemporally efficient extraction path follows the so-called Hotelling rule. The rate of increase in the price of a particular resource has to be equal to the real rate of interest. The intuition behind this result is easily outlined. The proceeds from extraction of a particular amount of a resource could be invested on capital markets at the given interest rate. Leaving the resource in the ground, by contrast, implies that the value of the resource stock has to increase by the same amount, implying that the price of every unit of the stock has to increase accordingly. Arbitrage ensures that the increase in the price in the case of non-extraction equals the interest rate as the measure of the return in the case of extraction (Olson and Knapp, 1997). As a consequence of these considerations, **prices** of non-renewable resources provide a direct signal for emerging problems of sustainability.

But economic theory reaches beyond this hypothetical ideal, emphasizing that "over-use" of non-renewable natural resources can occur as a consequence of **externalities** or of lacking **intergenerational fairness**. On the one hand, the extraction process itself can cause environmental, social and economic **damage** that is not reflected in the price of a particular resource. For example, a broad economic literature documents the so-called "resource curse": economies that are characterized by abundant supplies of a certain natural resource such as oil often experience slow growth, environmental degradation and social conflict.

On the other hand, economic efficiency alone might be an insufficient guide to sustainability with respect to the requirement that the current level of well-being should at least be maintained for future generations, owing to issues of intergenerational distribution and fairness. When taking into account the well-being of future generations, economically efficient allocations might be associated with undesired allocations of welfare between generations (Howarth, 1991). After all, unlike most other decisions that are taken today, the decision to extract non-renewable resources is genuinely **irreversible** (Sandler, 1997).

234. It therefore seems eminently sensible to augment the monitoring of current prices of non-renewable resource by indicators of their usage in production and consumption and their associated productivity. Using measures of resource productivity in conjunction with total known reserves may at first sight provide a rough estimate of total known future output given the current state of technology. We see two main **problems**, however, one conceptual in nature and the other regarding the appropriate interpretation of the measures provided. First, many conceptual issues arise in terms of **measurement**. In particular, there is not one homogeneous non-renewable resource but rather many and sometimes highly **heterogeneous** resources with varying degrees of substitutability. To ascertain parsimony, however, we seek to identify a small number of summary indicators.

Second, measuring resource productivity at a national or regional level cannot by itself provide an informative picture of the degree of sustainability at a **global** level. More concretely, measures of national resource productivity signal problems for sustainability only if one can establish a direct link between resource usage and domestic policy. Since such a measure necessarily mirrors **domestic** production patterns and mostly neglects the global perspective: a change in the indicator might signal **genuine changes** in resource productivity or the **outsourcing** of resource-intensive production elements abroad. Using the indicator in a mechani-

cal fashion might therefore lead to serious misconceptions regarding the development of resource usage.

Consequently, if accepted as a useful indicator, measures of domestic resource productivity would need to be supplemented by other indicators, as domestic resource productivity alone cannot comprehensively capture the resource usage of a nation or region (Box 5). The opportunities and benefits of globalization and international trade enable domestic consumption patterns to be decoupled from domestic production patterns. For example, one could introduce an additional indicator that does not focus on the production side of resource usage but rather on the **consumption** side. Domestic consumption includes not only domestic products that embody low resource usage due to highly efficient domestic production technology but also include imported products that include a huge amount of resources. A combination of the two indicators could, in principle, provide a better picture of the current path of domestic resource usage.

Box 5

Measures of resource productivity and consumption: current usage and problems

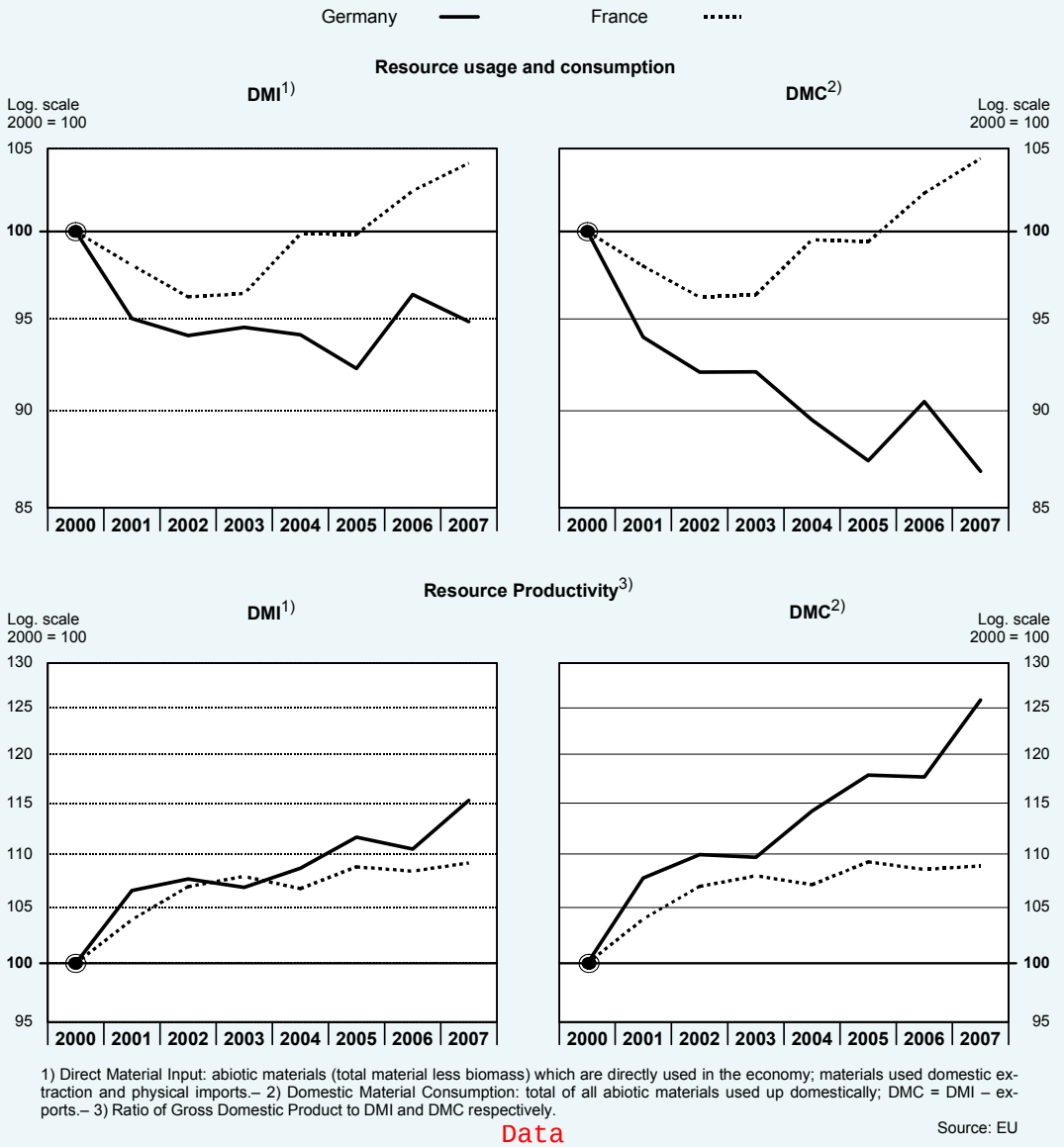
The indicator "resource productivity" is currently published by the German Federal Statistical Office at a national level and by Eurostat at a European level. However, the latter applies a slightly different definition as described below. The starting point for both published measures is **direct material input** (DMI), which aggregates the total amount (in tons) of primary resources that have been imported or extracted domestically and all finished and semi-finished imported goods. In the case of Germany this measure is limited to abiotic, i.e. non-renewable, resources. Subtracting exported primary resources and finished and semi-finished goods from DMI finally yields **domestic material consumption** (DMC).

Both measures pose a substantial number of methodological problems that can at best be attenuated, but not resolved completely. DMI measures usually lead to double counting internationally, as resources are measured on the exporter's and the importer's side. International comparisons are thus problematic, as aggregating national DMIs overstates the true global DMI. Consequently, national estimates can change simply because the direction of trade flows reverses, even though global resource productivity is unchanged. Focusing on the national perspective, the **evolution** of DMI and DMC in France and Germany between 2000 and 2007 is depicted in the top half of Chart 29. Whereas the DMI shows a similar pattern for both countries, the DMC show opposite patterns, indicating a decrease in material consumption in Germany and an increase in France between 2000 and 2007.

At the European level, the measure of **resource productivity** is defined as the ratio of real GDP to DMC. Germany uses a different definition, however. There resource productivity is defined as the ratio of real GDP to DMI. Strictly speaking, the German definition seems to fit the economic definition of productivity better than the European definition. GDP is an output measure, while DMC is a consumption measure, which makes the ratio difficult to interpret. The **evolution** of resource productivity in France and Germany between 2000 and 2007 is shown in the bottom half of Chart 29. Using DMI as a measure of resource usage shows very similar patterns for Germany and France. Both countries show a tendency towards increasing resource productivity. The consumption measure DMC provides a slightly different picture, while maintaining the qualitative result. Germany's resource productivity seems to have increased by more than 20 % while France shows an increase of 10 %.

Chart 29

Resource usage and consumption and resource productivity in Germany and France



1) Direct Material Input: abiotic materials (total material less biomass) which are directly used in the economy; materials used domestic extraction and physical imports.– 2) Domestic Material Consumption: total of all abiotic materials used up domestically; DMC = DMI – exports.– 3) Ratio of Gross Domestic Product to DMI and DMC respectively.

Data

Source: EU

An important drawback of this type of indicator is that it focuses primarily on sustainability in domestic terms without considering in detail the resource usage embodied in **imports**. This omission is becoming more severe over time, as structural and global change towards cross-border production networks with different stages of production scattered around the globe is leading to a relocation of resource use from one nation to another which is not captured in the current indicator. This shift towards the importation of (semi-)finished products embodying primary resources may even indicate a decrease in resource consumption, simply because primary imports of resources decrease as the domestic production of certain goods is substituted by the importation of such goods.

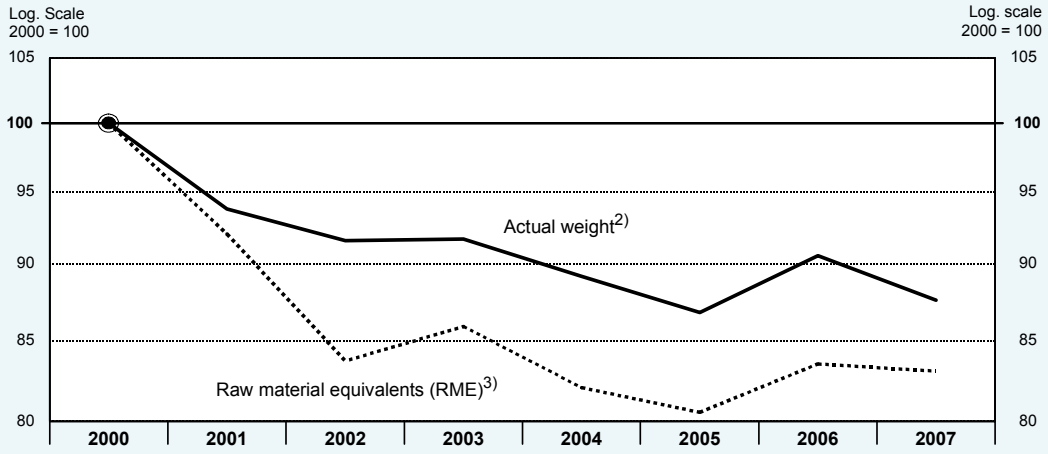
The DMC currently does not adequately reflect the resources embodied in imports and exports, since only their weight at the border is included, which in the case of final or semi-finished goods tends to **underestimate** their true resource content. Therefore, this indicator needs to be enhanced by a detailed calculation of imported and exported embodied resources to derive the global picture of domestic resource consumption. In the ideal case of correct measurement of

resource content it should then in principle be possible to calculate the world's "true" DMC by simply aggregating the DMC of all individual countries.

With a focus on the period 2000 to 2007, the German Federal Statistical Office attempted to adjust German figures for resource consumption by estimating in great detail the resource content of imported and exported (semi-finished) goods (Buyny and Lauber, 2010). This task was performed by breaking down the production process of many finished and semi-finished products with a view to assessing their resource usage and linking this with a comprehensive **input-output analysis** to end up with measures of DMI and DMC in raw material equivalents. The main findings of the research project are that, during the past decade, resource-intensive production has been shifted to other countries and that a significant share of German resource usage (DMI) is "concealed" in imported products and services. But most of the resources are further processed and eventually exported to other countries. Therefore, domestic consumption of resources (DMI less exports) in Germany actually decreased at a much higher rate (Chart 30) than that depicted in Chart 29.

Chart 30

Alternative measures of abiotic resource consumption (DMC) in Germany[1)]



1) DMC: total of all abiotic materials (total materials less biomass) used up domestically; DMC equals Direct Material Input (DMI) less exports.– 2) Values differ slightly due to other data source for abiotic resource usage.– 3) Instead of the total weight of imported goods, raw material equivalents represent the total direct and indirect resource input necessary in the production process abroad.

Data

Source: Destatis

However, even these adjusted figures still do not fully reflect imported resource usage as the input-output tables and processing technologies underlying those calculations are based on German input-output tables which are supplemented by generic production processes applied abroad. Considering that many sourcing countries may actually use technologies which are **not as efficient** as those applied by German firms, actual resource consumption may be even higher. Against this backdrop, it would be desirable to base these calculations on country-specific estimations of input-output tables and as well as processing technologies.

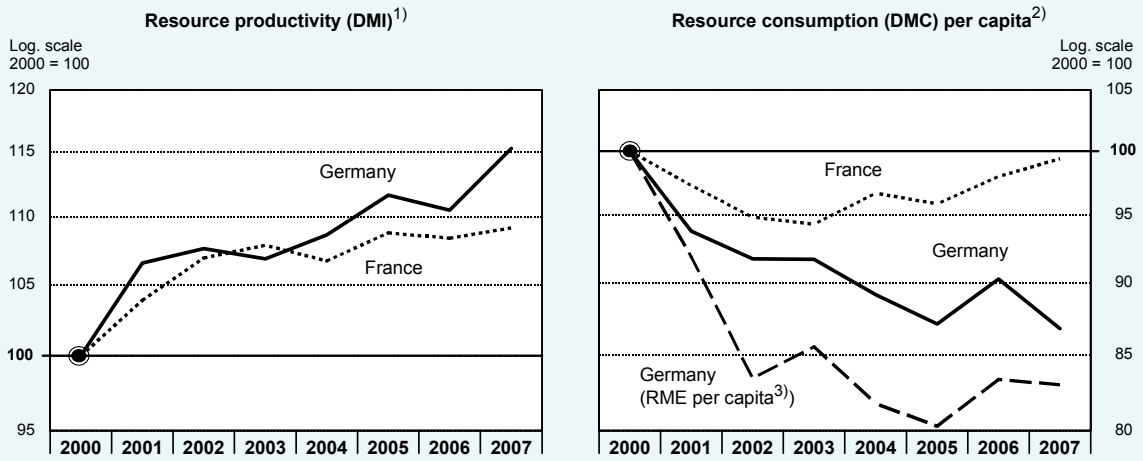
235. In sum, the conventional economic theory postulation that, barring externalities, all information relating to the sustainability of natural resources should be reflected in their prices can seriously be called into question. Whenever **prices** turn out to be **imperfect guides** to sustainability, the alternative of monitoring resource productivities might indeed be helpful in indicating unsustainable extraction paths. Since our discussion reveals, however, that we should be fairly sceptical as to whether indicators of resource productivity always provide

useful information on sustainability, we decided to include **two indicators** of **resource utilization** in the dashboard. As an indicator of resource productivity we report GDP relative to non-renewable DMI and as an indicator of resource consumption we report DMC per capita.

However, as argued above, the current measure of DMC as published at the European level and the national measure does not adequately reflect the resource consumption that is actually embodied in imports. Therefore, the measure of DMC – but not the productivity measure using DMI – should be further enhanced prospectively by calculating these in raw material equivalents such that domestic consumption is accounted for resources embodied in imports (Box 5). Both indicators of resource consumption are depicted in Chart 31, while for France

Chart 31

Resource productivity and resource consumption in Germany and France



1) Ratio of Gross Domestic Product to DMI; DMI (Direct Material Input): abiotic materials (total materials less biomass) which are directly used in the economy; materials used domestic extraction and physical imports).– 2) DMC: Domestic Material Consumption (total of all abiotic materials used up domestically; DMC = DMI – exports).– 3) DMC in raw material equivalents (RME: Instead of the total weight of imported goods, raw material equivalents represent the total direct and indirect resource input necessary in the production process abroad).

Data

Sources: Destatis, EU

no data is currently available on DMC in raw material equivalents. Despite our general reservations regarding these measures, this is certainly much better than starting from the assumption that current levels of resource extraction are too high and prescribing desired current depletion rates in relation to arbitrarily defined threshold levels. In the ideal case, the potential externalities associated with current modes of resource extraction would require a more elaborate monitoring system than that accomplished in our dashboard approach. We acknowledge that, owing to the heterogeneity of the issue and their imperfections, the entries in our dashboard can merely serve as warning signals.

Biodiversity

236. Biodiversity can be conceived as the totality of genes, species and ecosystems of a region and all their interactions. In principle, it can be seen as a form of **capital** that is needed to produce services needed to satisfy human needs. Arguably, its **preservation** is essential for food and nutrition security, medical progress, the chemical industry, industrial raw materials as well as ecosystem services like the absorption of carbon dioxide by oceans and forests (Baumgärtner, 2006). Correspondingly, declines and changes in biodiversity may negatively

impact the provision of food, fibre, medicine and fresh water as well as the pollination of crops, filtration of pollutants and protection from natural disasters.

It is possible to distinguish a **global** as well as a **local dimension** of biodiversity. For example, the absorption of carbon dioxide by oceans and forests is a global ecosystem service because most carbon dioxide is absorbed in locations different from the place in which the pollution originates. Conversely, ecosystem services that, for example, affect and improve soil fertility generate local biodiversity.

237. With a view to incorporating the potential importance of biodiversity services, a **summary indicator** of the stock of this type of natural capital should be incorporated in a comprehensive dashboard. To appreciate both the local and the global nature of the issue, a national indicator focusing on the ecosystems within a given territory should invariably be complemented by a figure representing the development of global biodiversity. Ideally, the selected indicators should be conceptually identical for national and for global use, to ensure comparability.

Two approaches to quantitatively assessing (changes in) biodiversity can be distinguished in the academic debate. While ecologists have traditionally employed concepts such as **species abundance**, economists, in turn, have tended to discuss concepts that are based on **pair-wise dissimilarity** between species or between weighted attributes of species, respectively. Obviously, these two classes of measures appreciate biodiversity for different reasons and assess aspects and components differently. As a consequence, the measurement of biodiversity requires **a normative judgement** as to what purpose biodiversity serves in ecological-economic systems (Baumgärtner, 2006).

238. There are many components of biodiversity dashboards, both **individual** and **composite indicators**, that are currently used to monitor the development of biodiversity. The German National Strategy of Biological Diversity, for example, uses indicators like the number of endangered species, the size of strictly protected areas, the increase in the amount of land used for human settlements and the transport infrastructure, organic farmland as a portion of total agricultural land, as well as the proportion of certified forest land. In our view, theses indicators are too selective, as they only mirror certain aspects of biodiversity. In particular, these indicators are not based on pair-wise dissimilarity, the concept economists prefer to use when discussing the issues of biodiversity.

The most widely used and most developed biodiversity indicators are **bird indices** as included in the German National Strategy of Biological Diversity, the German Strategy of Sustainable Development and the EU SDS. The German bird index refers to 59 bird species and specifies target values of the stock of each species in 2015. From an economic point of view, the logic of such target values remains ambiguous. Furthermore, there is no clear evidence that the number of birds is an appropriate proxy for biodiversity. Arguments for bird indices emphasize that birds can reflect changes in other dimensions of biodiversity and are responsive to environmental changes (Gregory et al., 2005). However, it is so far unclear what rela-

tionship exists between biodiversity captured in such a narrow way and the very broad range of potential services associated with complex ecosystems. Nevertheless, we decided to include this indicator as a **preliminary** entry in our dashboard.

239. Alternatively, some aggregate **biodiversity indices** have been developed, such as the Red List Index or the Living Planet Index (Cocciufa et al., 2006). Currently, the Living Planet Index is particularly appropriate on a **global scale**, but we need an indicator that is also appropriate on a national scale. Furthermore, the Living Planet Index is an aggregated indicator of species richness and it seems dubious whether such indicators are an appropriate proxy for biodiversity. Other aggregated indices, like the Red List Index, might fruitfully be applied to national levels, but unfortunately seem to be of rather **limited significance** to the issue at hand. Even if some species have a very high extinction risk at a national level, they may be classified as abundant at a global level.

Another approach to measuring biodiversity could be to account for biodiversity in **spatial planning processes**. In this approach, the first step would be a land cover database that is accurate, consistent over the national territory and regularly produced or updated. This database may be available at the European Level in the context of the CORINE (Coordinated Information on the European Environment) Land Cover Project. In a second step, the potential value of biodiversity of different land cover would have to be determined. In assigning biodiversity valuations to different usages, both the richness and rarity of species in different ecosystems would have to be assessed and the share of aboriginal or indigenous species relative to introduced species taken into account.

In France, a corresponding **pilot** scheme is underway in three départements to establish such a land cover database and introduce biodiversity monitoring. A similar pilot scheme was conducted in Germany between 1995 and 1996, called "ökologische Flächenstichprobe" (ecological area sample) (Hoffmann-Kroll et al., 1998). This sample was aimed at monitoring biodiversity and providing information about the state of landscapes. In the process, information about habitats, plant species, birds and other species was surveyed. It is important to stress that the "ökologische Flächenstichprobe" was not aimed at constructing a single indicator. However, it was established only in the state of North Rhine-Westphalia because decision makers considered the extension of this project to all of Germany as far too expensive. While the "ökologische Flächenstichprobe", as well as the French project, seem to be very promising projects that provide the data and tools necessary for capturing many aspects of biodiversity, they nonetheless cannot lay the basis for the type of regular statistical reporting that we are pursuing with our dashboard.

240. The importance of biodiversity and the potential impact resulting from an ongoing loss of biodiversity are undisputed. But all of the indicators discussed above were developed outside the realm of economics and so capture certain aspects of biodiversity that do not necessarily address issues which reflect the economically important aspects of biodiversity. Even though the year 2010 has been declared the International Year of Biodiversity by the United Nations, the state of **economic research** on biodiversity is still very much a **work in pro-**

gress. Therefore, economic biodiversity measures as proposed by Weitzmann (1992, 1993, 1998) or Nehring and Puppe (2002, 2004, 2009) are currently far from being operational and quantifiable. In 2007, the G8+5 Environmental Ministers launched a research initiative called "The Economics of Ecosystems and Biodiversity" (TEEB), but a first report to policy makers (TEEB, 2009) remains rather **vague** and does not provide any quantitative indicators which could enter our dashboard of sustainability indicators.

While we respect the important role that biodiversity plays in the context of environmental sustainability, we are not able to currently determine an explicit indicator that fully captures the economic dimension of biodiversity. Consequently, we decided to include the bird index in our dashboard as the **preliminary** fifth entry regarding environmental sustainability. Having said that, we emphatically recommend conducting further research on this issue in the hope that this will lead to a more appropriate entry in the dashboard in the coming years.

5. Concluding remarks

241. This chapter began from the insight that, although current economic performance and well-being might be quite satisfactory at the moment, present paths of action, if persistently continued into the future, might well turn out to be unsustainable. In that case, they might require sharp and painful adjustments and perhaps even cause socially costly crises. One section of this chapter focused on two facets of **economic sustainability** – growth sustainability, on the one hand, and external and fiscal sustainability, on the other. Another section addressed a third facet, namely private sector financial sustainability. Throughout these sections the discussion combined the **medium-term** and the **long-term** perspective, since in this context the well-being of future generations tends to be closely related to what happens to the current generation in the medium run.

242. The first aspect of economic sustainability which we have addressed in our analysis is the issue of **growth sustainability**. Specifically, we consider growth to be sustainable if a sufficient part of wealth creation in the economy is allocated to investment, irrespective of whether it is invested in material or immaterial capacities. Consequently, in order to emphasize the importance of capital accumulation for economic growth, we have decided to include the ratio of **private sector net fixed capital formation** to GDP in our dashboard. Moreover, since we require a reliable predictor of future overall productivity and of expected trends in science, technology and innovation, we have chosen as a second indicator of growth sustainability for our dashboard **R&D investment** of an economy relative to its GDP.

243. The second aspect of economic sustainability, external and **fiscal sustainability**, is integrally related to the intertemporal budget constraint which is necessarily binding in the long term. Due to its inherent long-term perspective, this issue is also closely linked to concerns of intergenerational equity. When unsustainable fiscal and external positions ultimately have to be unwound, this can have painful consequences. As our concrete indicators of fiscal sustainability, we have chosen, first, the **cyclically-adjusted public sector balance**, which, according to the Golden Rule of Public Finance, should not exceed net public investment. And as a second indicator of fiscal sustainability we have selected the **fiscal sustainability gap** as rep-

resented by the indicator "S2" in the European Commission's Sustainability Reports. To signal fiscal sustainability, this indicator should not be positive. Given a positive sustainability gap, the indicator should at least decrease over time and eventually converge towards zero if recipients are to rest assured that current fiscal policies are sustainable.

244. The present chapter has also discussed possibilities of augmenting the monitoring of current economic performance and well-being that is regularly conducted by the statistical offices by a complementary documentation of the state of financial sustainability. To this end, it suggests a **set of indicators** which signal unsustainable developments in the private and in the financial sector. Their objective is exclusively to investigate **excessive** fundamental and undesirable developments which are likely to lead to severe economic crises. While this objective is ambitious, the discussion has made clear that it will never be possible to predict financial crises with certainty. What is offered here, though, is a small set of reasonably robust **early-warning** indicators which could alert policy makers and the general public in the event of fundamental undesirable developments in the financial sector. They are intended to be simple and manageable indicators for policy makers and the wider public who themselves do not have the time and expertise to consider a plethora of disaggregated indicators or to employ stress testing or early-warning models.

245. Despite these reservations, the three proposed indicators are – in our opinion – the most reasonable extract of empirical literature concerned with the issue of leading indicators. Specifically, we suggest looking at the ratio of total **private credit** relative to GDP, and at **real equity prices** as well as **real property prices**, both deflated by the consumer price index. This suggestion can be implemented directly. Data on total private credit and equity prices are provided by national central banks, and data on property prices are collected by the Bank of International Settlements (BIS) and can be retrieved from the BIS (BIS, 2010).While this limited set of indicators should obviously not be understood as a substitute for detailed macroprudential supervision or existing early-warning systems used by experts and sovereign authorities, their promise is to identify early in the process those economic developments which, if left uncorrected, might lead to distressed situations. If these indicators signal an alarming development, policy makers should consult experts and authorities and possibly take remedial action.

Regarding further work on this issue, especially on the supranational level, it is necessary to ensure data quality. For instance, Borio and Drehmann argue that data are still subject to limitations such as heterogeneity across countries (Borio and Drehmann, 2009a; McKinsey, 2010). Hence, there is a need for **harmonization** and **standardization** of data collection processes across countries in order to generate reliable and comparable information. This is all the more important as globalization in general and financial integration in particular force us to act at the EU-level – thereby involving 27 nation states. As harmonization is primarily a matter of setting standards for definitions, data collection processes and data quality, this should be a very cost-efficient but simultaneously valuable contribution.

246. According to the current state of knowledge, rising levels of carbon dioxide and other greenhouse gases in the atmosphere have already caused global warming, and will induce climate change on an even broader scale going forward. Climate change has the potential to trigger major social and economic crises. Accordingly, GHG emissions should be a component of our dashboard. Of course, the figure which is most relevant for climate change is the **level** of **GHG emissions**. But climate change is a global phenomenon and because of this the national indicator of GHG emissions, expressed in level terms, which we propose for our dashboard, could be highly misleading when considered in isolation. Thus, in a dashboard it should always be complemented by some summary figures documenting total GHG emissions or, in lieu of complete data, CO_2 emissions.

Obviously, an appropriate strategy limiting global anthropogenic GHG emissions requires a binding international agreement. Key elements of such an agreement should be a legally binding target of greenhouse gas emissions, an international emission trading system and an allocation mechanism that distributes emission allowances among the participating countries. Even though quite different allocation mechanisms are conceivable, the principle of equality seems to be a good starting point for a fair distribution of the global budget. Therefore, equal per-capita emission rights all over the world would arguably form a sensible basis for the allocation of national emissions budgets. Irrespective of its potential role in an allocation mechanism for globally traded emission permits, it would be advisable to inform policy makers and the general public about national **GHG emissions per capita**. Hence, we propose to include the current GHG emissions per capita as a second GHG indicator in our dashboard.

247. The sustainability of (non-renewable) resources has been a hotly debated topic for decades among policy makers, scientists and the wider public alike. From the vantage point of economic theory, an emerging scarcity of non-renewable resources is primarily reflected in the evolution of their prices, and additional monitoring of physical measures does not seem necessary. But economic theory reaches beyond this hypothetical ideal, emphasizing the potential "over-use" of non-renewable natural resources that can occur as a consequence of **externalities** or of lacking **intergenerational fairness**. Therefore, beyond current prices we propose monitoring physical flows of non-renewable resource. This can be achieved by publishing indicators of non-renewable resource usage in production and consumption and their associated productivity, i.e. GDP relative to these measures. Our proposed first measure is direct material input (DMI) which comprises the total amount of raw non-renewable resources used in domestic production. Our proposed second measure is domestic material consumption (DMC) per capita, which measures the amount of resources consumed domestically by deducting exports from DMI. Prospectively, DMC should be enhanced to account for the resource content of imported and exported goods.

248. At least in a narrow sense, biodiversity is a form of capital that is required to produce services intended to satisfy human needs. Arguably, its preservation is essential for many desirable facets of current and future human existence, like food and nutrition security, medical progress or industrial raw materials. Moreover, ascertaining biodiversity is not only a global issue, but also relates to the stability of local ecosystems. Given its importance, a **biodiversity**

indicator should therefore be added to our dashboard. Unfortunately, all existing indicators were developed outside the realm of economics, making it difficult to gauge whether they fully account for possible welfare trade-offs involved within and across generations. While we are unable to currently determine an explicit indicator that fully captures the economic dimension of biodiversity, we decided to include the bird index as the preliminary fifth entry in our dashboard regarding environmental sustainability.

Appendix: Financial Sustainability

Meaningful leading indicators have to satisfy requirements along various dimensions. This appendix provides some guidelines on the way to selecting reliable and robust leading indicators from a set of possible candidates presented in Table A1.

First, it is important that leading indicators have the general ability to forewarn authorities of impending crises and thus have an appropriate **out-of-sample performance** (Borio and Drehmann, 2009a; Davis and Karim, 2008a). Many of the proposed leading indicators perform quite well in-sample. However, as already noted above, it is necessary to determine a set of indicators that provide a good summary or overview over a broad range of financial developments so as to cover the perimeter as far as possible. This is especially important as the particular causes and roots typically differ from crisis to crisis (Ghosh et al., 2009).

Second, to be useful for economic policy, any indicator has to identify the risk of future financial strains with a **lead sufficient** to allow the authorities to take remedial action (Borio and Drehmann, 2009a). Using the notion of Borio and Drehmann, this applies to the question of how far indicators act as barometers rather than thermometers of financial distress (Borio and Drehmann, 2009b). While earlier studies considered **forecasting horizons** of about one year (Kaminsky and Reinhart, 1999), later studies use longer and multiple horizons (Borio and Lowe, 2002a). Currently, forecasting horizons range from one to four years ahead, giving policy makers time to consider the situation and implement appropriate measures (Borio and Drehmann, 2009b). Naturally, from the vantage point of the policy maker, the earlier signal is the more valuable one (Kaminsky and Reinhart, 1999).

Third, another crucial aspect concerning leading indicators is the **availability of data**. The primary paradigm should be the selection of data with a sufficient lead and confidence (Borio and Drehmann, 2009b). Considerable data limitations typically exist in terms of **quality** and **quantity**. For instance, data collection is often not standardized across countries so that country-heterogeneity is an inherent problem, especially for cross-country investigations (McKinsey, 2010). This might have considerable consequences as a number of authors use data availability as a criterion for selecting which countries and data to include in the study, which in itself could give rise to sample selection bias (Bell and Pain, 2000). In addition, data should be available in a **timely** manner and reporting lags should be avoided, i.e. timely updates of the leading indicators must be available. Data which are available only with a substantial delay considerably shorten the forecasting horizon and the predictive ability of the leading indicators.

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
Borio and Drehmann, (2009a)	– crisis prediction performance of indicators using thresholds – composite index	– equity prices – private credit/GDP – property prices – cross-country exposures	1970-2008, 18 industrial countries	– Study is based on (Borio and Lowe, 2002a, b). – Indicators based exclusively on credit and equity prices would have failed to issue warning signals for the current crisis (out-of-sample over the period 2004-2008). – Performance of the indicator that additionally includes (commercial and residential) property prices improves the prediction performance substantially. – Additionally taking into account cross-country exposures has the potential to improve the prediction performance even further. – Future research: taking account of credit risk spreads and leverage.
Misina and Tkacz, (2009)	– linear and non-linear models	– credit measures (total domestic business credit) – asset prices (commercial real estate index, residential real estate index)	1984-2006, Canada, Japan, USA	– As it is difficult to test various leading indicators for individual developed countries that have faced very few, if any, financial crises in the past, the study tries to circumvent this problem by focusing on financial stress using the financial stress index (FSI) developed by (Illing and Liu, 2006)). – The exercise is mainly performed for Canada, but the robustness checks also consider Japan and the USA. – Within a linear framework, domestic credit growth is the best predictor of the financial stress index (FSI) at all horizons. – Asset prices tend to be better predictors of stress in a non-linear model. – At a two-year horizon, business credit and real estate prices (commercial real estate and new house prices) emerge as important predictors of financial stress. – The results summarize the performance of out-of-sample predictions.
Rose and Spiegel, (2009)	– multiple-indicator multiple cause (MIMIC) econometric model	– change in stock market capitalization/GDP – current account/GDP – short-term debt/reserves – domestic bank credit/GDP – bank claims/deposits	2008, 107 countries	– The study is able to reasonably model the severity of the crisis across countries, but it is unable to link it empirically to country-specific causes. – There are a few exceptions to the generally weak results: – Countries that experienced a large bull run in their stock market (measured relative to output) between 2003 and 2006 were more likely to be hit by the 2008 crisis. – Countries with larger current account deficits and fewer reserves (measured relative to short-term debt) were also more vulnerable. – There is weaker evidence that countries with high credit growth and a more levered banking sector are also associated with the severity of the crisis. – Some of the Eastern European and Baltic countries have been hard-hit,

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
				and this is apparent when we include geographic dummies.
Schularick and Taylor, (2009)	– probabilistic (OLS linear probability and logit) model	– real aggregate bank loans	1870-2008, 12 developed countries	– Lagged credit growth (defined as real aggregate bank loans) turns out to be highly significant as a predictor of financial crises. – Robustness tests show that past growth of credit emerges as the single best predictor of future financial instability. – The model also proved informative out-of-sample at the 5 per cent significance level. – Drawbacks: definitions of credit, money and banking institutions vary profoundly across countries, which makes cross-country comparisons difficult.
Davis and Karim, (2008a)	– crisis prediction performance of indicators using thresholds – composite index – multivariate logit model	– real GDP growth – real interest rate – real GDP per capita – change in terms of trade – fiscal balance/GDP – M2/reserves – private credit/GDP – credit growth – deposit insurance	1979-2003, 105 countries	– In the basis regression model, real GDP growth, real interest rates, real GDP per capita and terms-of-trade changes are consistently and significantly associated with crises. – Transforming independent variables (including standardization, lags, and interaction terms) improves the model with further variables becoming significant: fiscal balance/GDP, M2/reserves, private credit/GDP, credit growth and deposit insurance schemes. – Combining variables into composite indicators improves crisis prediction ability. – Out-of-sample results suggest that the model would provide valuable information for policy makers.
Davis and Karim, (2008b)	– multvariate logit model – binary recursive tree (BRT) approach	logit indicators: – real GDP growth – terms of trade – real GDP per capita – M2/reserves BRT indicators: – real domestic credit growth – real interest rate – nominal exchange rate – inflation	1979-2007, 7 advanced OECD countries and 65 emerging markets	– The paper uses two approaches to predict out-of-sample banking crises. – The logit approach determines real GDP growth, terms of trade, real GDP per capita and M2/reserves to be significant indicators of banking crises. – The BRT model identifies indicators of banking crises in the following order (according to their importance): real domestic credit growth, real interest rates, nominal exchange rate, inflation.
Duttagupta and Cashin, (2008)	– binary recursive tree (BRT) approach	– nominal exchange rate – interest rate margin – inflation – foreign deposits/foreign reserves – private credit/deposits	1990-2005, 50 emerging market and developing countries	– The BRT identifies five variables as the most important determinants of banking crises: nominal depreciation, bank profitability (interest rate margin), inflation, liability dollarization (foreign deposits/reserves), and bank liquidity (credit/deposits). – It also identifies three key crisis-prone conditions: – Macroeconomic instability: high annual inflation combined with relatively low terms-of-trade growth. – Low bank profitability: low interest profitability (spread between lending

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
				and deposit rates) combined with modest export growth.
				– High foreign exchange risk: high liability dollarization (foreign deposits/foreign reserves) combined with either (i) relatively high depreciation or (ii) low bank liquidity (private credit/deposits).
Hanschel and Monnin, (2005)	– multvariate model	– GDP – European GDP – asset prices (share prices and housing prices) – credit to private sector/GDP – investment/GDP	1987-2002, Switzerland	– The model uses a stress index, summarizing the banking sector's condition, as the dependent variable. – Independent variables are expressed as deviations from trend. – The model predicts the major stress periods. – The model performs well also for out-of-sample forecasts. – In various specifications, three variables appear to be robust: equity prices, housing prices, credit ratio. – The lag between the indicators and the index which represents stress periods could extend from one up to five years.
Noy, (2004)	– multivariate probit model	– domestic financial liberalization – inflation rate – M2/reserves – GDP per capita growth – foreign interest rate – real exchange rate	1975-1997, 61 non-OECD countries	– An increase in the inflation rate, an increase in the M2/reserves ratio, a decrease in GDP growth rate, a depreciating real exchange rate, and a decrease in foreign interest rates are all found to contribute to the likelihood of a banking crisis, as theory suggests. – Financial liberalization is a significant indicator of banking crises.
Borio and Lowe, (2002a)	– crisis prediction performance of indicators using thresholds – composite index	– real asset prices (real equity prices) – total private credit/GDP – investment	1960-1999, 34 emerging and industrial countries (including G10)	– Sustained rapid credit growth combined with large increases in asset prices appears to increase the probability of financial instability. Adding investment makes no significant improvement. – It is the combination of indicators that matters, not single indicators alone. I.e. Interactions between various imbalances is particularly important. – The relevant issue is not whether a "bubble" exists in a given asset price, but rather what combination of events in the financial and real sectors exposes the financial system to a materially increased level of risk.
Borio and Lowe, (2002b)	– crisis prediction performance of indicators using thresholds – composite index	– real asset prices (real equity prices) – private credit/GDP – real effective exchange rate	1960-1999, 34 emerging and industrial countries	– Credit, equity price and exchange rate contain useful joint information (all country setting). – For industrial countries, the inclusion of exchange rates does not improve predictive performance significantly. – For emerging markets, the exchange rate adds useful information.

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
Eichengreen and Arteta, (2000)	– multivariate probit regression	– domestic credit growth – M2/reserves – budget balance/GNP – current account/GDP	1975-1997, 75 emerging markets	– Domestic credit booms are strongly associated with banking crises. – Low reserves (proxied by M2) may be another symptom of rapid credit growth that sets the stage for crises. – Budget surpluses rather than deficits are associated with banking crises. – Current account balance/GDP ratio is actually significant in many regressions.
Kaminsky, (2000)	– signaling approach with composite leading indicators	– M2 multiplier – domestic credit/GDP – financial liberalization – excess M1 balances – exports – terms of trade – real exchange rate – reserves – M2/reserves – real interest rate differential – global real interest rate – foreign debt – capital flight – short-term foreign debt – output – domestic real interest rate – stock prices	1970-1995, 20 countries	– Note that only leading indicators with a noise-to-signal-ratio (NSR) of less than 1.0 are listed here. – Using the leading indicators, various composite indicators are constructed to estimate the probabilities of impending crises. – According to the composite indicators, the average probability of banking problems increases from 8 per cent in tranquil times to 17 per cent in times of distress, suggesting an increase in the vulnerability of the economy in the midst of banking problems. – In sum, all composite indicators perform better in predicting crises compared to the single best indicator, namely real exchange rate.
Demirgüc-Kunt and Detragiache, (1999)	– multivariate logit model	– real GDP growth – real interest rate – inflation – M2/reserves – credit growth	1980-1995, 65 developing and developed countries	– The study is closely related to previous works by Demirgüc-Kunt and Detragiache and is mainly focused on the elements for a “ready-to-use” procedure for decision makers. – Low GDP growth, a high real interest rate, high inflation, strong growth of bank credit in the past and a large ratio of broad money to reserves are all associated with a high probability of a banking crisis.
Gonzalez-Hermosillo, (1999)	– multivariate logit model	– nonperforming loans/total assets – capital/total assets – commercial and industrial loans/total assets – agricultural production loans/total assets – construction loans plus loans secured by multifamily, nonresidential, and farm real estate/total assets – loans secured by	1980-1995, U.S. Southwest, U.S. Northeast, California, Mexico, Columbia,	– A high ratio of nonperforming loans to total assets and a low ratio of capital to total assets increase the probability of bank failure and distress. – The proxies for market risk (commercial and industrial loans, agricultural production loans, construction loans plus loans secured by multifamily, nonresidential and farm real estate, loans secured by family real estate, housing loans, consumer loans, unsecuritized loans) and liquidity risk (large certificates of deposit, deposits from the public, Fed funds purchased plus other borrowed funds, deposits from other banks, investment securities, interest expenditures) were generally important

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
		family real estate/total assets – housing loans/total assets – consumer loans/total assets – unsecuritized loans/total assets – large certificates of deposit/total assets – deposits from the public/total assets		in determining bank distress and eventual failure.
		– Fed funds purchased plus other borrowed funds/total assets – deposits from other banks/total assets – investment securities/total assets – interest expenditures/total deposits		
Kaminsky and Reinhart, (1999)	– signaling approach	– M2 multiplier – domestic credit/GDP – real interest rate – M2/reserves – exports – real exchange rate – imports – reserves – real interest-rate differential – output – stock prices	1970-1995, 20 developing and developed countries	– Note that only indicators which i) called at least 50 per cent of crises accurately and ii) with a noise-to-signal-ratio of less than 1.0 are listed here. – The forecasting window is 12 months. – Each of the following indicators signaled more than 80 per cent of crises: real interest rate, exports, reserves, real interest-rate differential, output and stock prices. – Each of the following indicators has a noise-to-signal-ratio of 50 per cent or less: M2 multiplier, real interest rate, imports, output and stock prices. – Some general conclusion apply: – Banking crises are preceded by recessions or, at least, below-average economic growth. – The financial vulnerability of the economy increases as the unsecuritized liabilities of the banking-system climb to lofty levels. – Crises are typically preceded by a multitude of weak and deteriorating economic fundamentals.
Hardy and Pazarbasioglu, (1998)	– multivariate logit model	– capital output ratio – inflation – real interest rates – real effective exchange rate – gross foreign liabilities/GDP – terms of trade	1980-1997, 38 countries	– The best warning signs were proxies for the vulnerability of the banking and corporate sector, such as credit growth and rising foreign liabilities. – Capital-output ratio is not significant, but including the variables improves predictive power. – A rise followed by a sharp fall in inflation seems to be one of the most reliable early indicators of impending banking sector problems. – Real interest rates usually rise in the crisis year, and reliably tend to start increasing already in the preceding years. – Banking crises are associated with a sharp decline in the real effective ex-

Table A1: Empirical literature on early warning systems for banking crises

Study	Methodology	Significant and robust core indicators	Sample coverage	Core findings
				change rate (REER). However, an appreciation of the REER often precedes a crisis. – Gross foreign liabilities of the banking sector relative to GDP are significant and contribute to the predictive power of the model. – A decrease in terms of trade is found to significantly precede a crisis situation. – The inclusion of regional variables improves the predictive power of the model.
				– About one third of crises can be predicted using just leading indicators
Demirgüc-Kunt and Detragiache, (1997)	– multivariate logit model	– real GDP growth – inflation – real interest rates – M2/foreign exchange reserves – domestic credit to the private sector/GDP – real domestic credit growth – deposit insurance schemes – law & order index	1980-1994, 65 developed and developing countries	– Crises tend to erupt in a weak macroeconomic environment characterized by slow GDP growth and high inflation. GDP growth loses significance if it is lagged by one period. – High real interest rates are clearly associated with systemic banking sector problems. – There is some evidence that vulnerability to balance of payments crises has played a role: the tests also indicate that vulnerability to sudden capital outflows, a high share of credit to the private sector (less robust), and high past credit growth (less robust) may be associated with a higher probability of a crisis. – Countries with explicit deposit insurance schemes were particularly at risk, as were countries with weak law enforcement.

References

Baumgärtner, S. (2006) *Measuring the Diversity of What? And for What Purpose? A Conceptual Comparison of Ecological and Economic Biodiversity Indices*, http://ssrn.com/abstract=894782 (accessed 5.7.2010).

Bell, J. and D. Pain (2000) *Leading Indicator Models of Banking Crises – A Critical Review*, Financial Stability Review.

BIS (2010) *Property Price Statistics*, Bank for International Settlements.

Blanchard, O. J., J. Chouraqui, R. P. Hagemann and N. Sartor (1990) *The Sustainability of Fiscal Policy: New Answers to an Old Question*, OECD Economic Studies, 15, 7 - 34.

Bordo, M., B. Eichengreen, D. Klingenbiel and M. S. Martinez-Peria (2001) *Is the Crisis Problem Growing more Severe?*, Economic Policy, 16 (32), 51 - 82.

Borio, C. and M. Drehmann (2009a) *Assessing the Risk of Banking Crises – Revisited*, BIS Quarterly Review, March, 29 - 46.

--- (2009b) *Towards an Operational Framework for Financial Stability: "Fuzzy" Measurement and its Consequences*, Bank for International Settlements, 284.

Borio, C. and P. Lowe (2002a) *Asset Prices, Financial and Monetary Stability: Exploring the Nexus*, Bank for International Settlements, 114.

--- (2002b) *Assessing the Risk of Banking Crises*, BIS Quarterly Review, 43 - 54.

Buyny, Š. and U. Lauber (2010) *Calculating Imports and Exports in Raw Material Equivalents*, Environmental-Economic Accounting (EEA), Federal Statistical Office Germany.

Caprio, G. and D. Klingebiel (1996) *Bank Insolvencies: Cross-Country Experience*, World Bank Policy Research Working Paper No. 1620.

Cocciufa, C., M. Bredemeier, L. Halada, C. Klok, B. Normander, B. Petriccione and R. Smulders (2006) *State of European Biodiversity Completed Projects - Expertise Centre for Biodiversity and Sustainable Development*, http://www.ecnc.org/programmes/state-of-european-biodiversity-completed-projects?action=detail&id=76 (accessed 2010-06-29).

Dasgupta, P. and G. Heal (1974) *The Optimal Depletion of Exhaustible Resources*, The Review of Economic Studies, 41, 3 - 28.

d'Autume, A. and K. Schubert (2008) *Hartwick's Rule and Maximin Paths when the Exhaustible Resource has an Amenity Value*, Journal of Environmental Economics and Management, 56 (3), 260 - 274.

Davis, P. and D. Karim (2008a) *Comparing Early Warning Systems for Banking Crises*, Journal of Financial Stability, 4, 89 - 120.

--- (2008b) *Could Early Warning Systems Have Helped to Predict the Sub Prime Crisis?*, National Institute Economic Review, 206 (1), 25 - 37.

Demirgüc-Kunt, A. and E. Detragiache (1997) *The Determinants of Banking Crises: Evidence from Developing and Developed Countries*, IMF Working Paper No. 97/106.

--- (1999) *Monitoring Banking Sector Fragility: A Multivariate Logit Approach*, IMF Working Paper No. 99/147.

Duttagupta, R. and P. Cashin (2008) *The Anatomy of Banking Crises*, IMF Working Paper No. 08/93.

Eichengreen, B. and C. Arteta (2000) *Banking Crises in Emerging Markets: Presumptions and Evidence*, CIDER Working Paper No. 115, Centre for International and Development Economics Research, University of California.

European Commission (2007) *Progress Report of the Sustainability Development Strategy 2007*.

--- (2010) *EUROPE 2020 A Strategy for Smart, Sustainable and Inclusive Growth*, Communications of the European Commission.

Eurostat (2007) *Measuring Progress towards a more Sustainable Europe - 2007 Monitoring Report of the EU Sustainable Development Strategy*.

Frascati Manual (2002) *The Measurement of Scientific and Technological Activities - Proposed Standard Practice for Surveys on Research and Experimental Development*, OECD, Paris.

Ghosh, A. R., J. D. Ostry and N. Tamirisa (2009) *Anticipating the Next Crisis*, Finance and Development, 46 (3).

Gonzalez-Hermosillo, B. (1999) *Determinants of Ex-Ante Banking System Distress: A Macro-Micro Empirical Exploration of Some Recent Episodes*, IMF Working Paper No. 33/99.

Gregory, R. D., R. P. Foppen, D. W. Gibbons, A. W. Gmelig Meyling, D. G. Noble, A. van Strien and P. Vorisek (2005) *Developing Indicators for European Birds*, Philosophical Transactions of the Royal Society B: Biological Sciences, 360 (1454), 269 - 288.

Griliches, Z. (1986) *Productivity, R&D, and the Basic Research at the Firm Level in the 1970's*, American Economic Review, 76 (1), 141 - 154.

Griliches, Z. and F. R. Lichtenberg (1982) *R&D and Productivity at the Industry Level: Is There Still a Relationship?*, NBER Working Paper, No. 850, National Bureau of Economic Research.

Hamilton, K. and M. Clemens (1999) *Genuine Savings Rates in Developing Countries*, World Bank Economic Review, 13 (2), 333 - 356.

Hanschel, E. and P. Monnin (2005) *Measuring and Forecasting Stress in the Banking Sector: Evidence from Switzerland*, BIS Papers No. 22.

Hardy, D. and C. Pazarbasioglu (1998) *Leading Indicators of Banking Crises: Was Asia Different?*, IMF Working Paper No. 98/91.

Harris, J., T. Wise, K. Gallagher, N. R. Goodwin and A. Sen (2001) *A Survey of Sustainable Development: Social And Economic Dimensions*, 1. ed., Island Press.

Hartwick, J. M. (1977) *Intergenerational Equity and the Investing of Rents from Exhaustible Resources*, American Economic Review, 67 (5), 972 - 974.

Hoffmann-Kroll, R., D. Schäfer and S. Seibel (1998) *Biodiversität und Statistik. Ergebnisse des Pilotprojekts zur ökologischen Flächenstichprobe*, Wirtschaft und Statistik (1), 60 - 75.

Hoggarth, G., R. Reis and V. Saporta (2001) *Costs of Banking System Instability: Some Empirical Evidence*, Bank of England Working Paper No. 144.

Horton, M. and A. Ivanova (2009) *The Size of Fiscal Expansion: An Analysis for the Largest Countries*, IMF Technical Report.

Howarth, R. B. (1991) *Intertemporal Equilibria and Exhaustible Resources: An Overlapping Generations Approach*, Ecological Economics, 4 (3), 237 - 252.

Howitt, P. (2000) *Endogenous Growth and Cross-country Income Differences*, American Economic Review, 90 (4), 829 - 846.

Illing, M. and Y. Liu (2006) *Measuring Financial Stress in a Developed Country: An Application to Canada*, Journal of Financial Stability, 2 (3), 243 - 265.

IMF (2009) *Global Financial Stability Report*, International Monetary Fund, 44a.

Jones, C. I. (2002) *Sources of U.S. Economic Growth in a World of Ideas*, American Economic Review, 92 (1), 220 - 239.

Kaminsky, G. (2000) *Currency and Banking Crises: The Early Warnings of Distress*, George Washington University.

Kaminsky, G. and C. Reinhart (1999) *The Twin Crises: The Causes of Banking and Balance-of-Payments Problems*, American Economic Review, 89 (3), 473 - 500.

Kindleberger, C. (1978) *Manias, Panics, and Crashes*, New York: Basic Books.

Laeven, L. and F. Valencia (2008) *Systemic Banking Crises: A New Database*, IMF Working Paper No. 08/224.

McKinsey (2010) *Debt and Deleveraging: The Global Credit Bubble and its Economic Consequences*, McKinsey&Company.

Minsky, H. (2008) *Stabilizing an Unstable Economy*, 2. ed., New York: Mc Graw Hill.

Misina, M. and G. Tkacz (2009) *Credit, Asset Prices, and Financial Stress*, International Journal of Central Banking, 5 (4), 95 - 122.

Nehring, K. and C. Puppe (2002) *A Theory of Diversity*, Econometrica (70), 1155 - 1198.

--- (2004) *Modelling phylogenetic diversity*, Resource and Energy Economics (26), 205 - 235.

--- (2009) *Diversity*, in: Anand, P., P. K. Pattanaik and C. Puppe (eds.): Handbook of rational & social choice, Oxford University Press.

Noy, I. (2004) *Financial Liberalization, Prudential Supervision, and the Onset of Banking Crises*, Emerging Markets Review, 5, 341 - 359.

Olson, L. J. and K. C. Knapp (1997) *Exhaustible Resource Allocation in an Overlapping Generations Economy*, Journal of Environmental Economics and Management, 32 (3), 277 - 292.

Pearce, D., K. Hamilton and G. Atkinson (1996) *Measuring Sustainable Development: Progress on Indicators*, Environment and Development Economics, 1 (01), 85 - 101.

Polasky, S., C. Costello and A. Solow (2005) *The Economics of Biodiversity*, in: Mäler, Karl-Göran and Jeffrey R. Vincent (eds.): Economywide and International Environmental Issues Handbook of Environmental Economics, vol. 3, Elsevier, 1517 - 1560.

Reinhart, C. and K. Rogoff (2009) *The Aftermath of Financial Crises*, American Economic Review, 99, 466 - 472.

--- (2010a) *Growth in a Time of Debt* (100), 573 - 578.

--- (2010b) *From Financial Crash to Debt Crisis*, NBER Working Paper No. 15795.

Romer, P. M. (1990) *Endogenous technological change*, Journal of Political Economy, 98 (5), 71 - 102.

Rose, A. and M. Spiegel (2009) *Cross-Country Causes and Consequences of the 2008 Crisis: Early Warning*, Federal Reserve Bank of San Francisco, Working Paper Series, Working Paper No. 2009-17.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2007) *Staatsverschuldung wirksam begrenzen*, Expertise im Auftrag des Bundesministeriums für Wirtschaft und Technologie, Wiesbaden.

--- (2009) *Deutschland im internationalen Konjunkturzusammenhang*, Expertise im Auftrag des Bundesministeriums für Wirtschaft und Technologie, Wiesbaden.

Saint-Etienne, C. (2004) *Une réforme politiquement acceptable du Pacte de stabilité et de croissance*, Conseil d'Analyse Economique, La documentation française.

Sandler, T. (1997) *Global Challenges*, Cambridge University Press.

Schularick, M. and A.M. Taylor (2009) *Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870-2008*, Working Paper, No. 15512, National Bureau of Economic Research.

Solow, R. M. (1974) *Intergenerational Equity and Exhaustible Resources*, The Review of Economic Studies, 41, 29 - 45.

Stern, N. H. (2007) *The Economics of Climate Change: The Stern Review*, Cambridge, UK: Cambridge University Press.

TEEB (2009) *The Economics of Ecosystems and Biodiversity for National and International Policy Makers – Summary: Responding to the Value of Nature 2009*.

Tirole, J. (2009) *Climate Policy: A New International Architecture*, Analyses Économiques, The Newsletter of the French Council of Economic Analysis, Vol. VII-06 (Oct. 2009).

Weitzmann, M. (1992) *What to Preserve? An Application of Diversity Theory to Crane Conservation*, Quarterly Journal of Economics (108), 155 - 183.

--- (1993) *On Diversity*, Quarterly Journal of Economics (107), 363 - 405.

--- (1998) *The Noah's Ark Problem*, Econometrica (66), 1279 - 1298.

Wissenschaftlicher Beirat beim Bundesministerium der Finanzen (2010) *Klimapolitik zwischen Emissionsvermeidung und Anpassung*.

Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (2009a) *Factsheet 2/2009 Climate change: Why 2°C?*

--- (2009b) *Solving the Climate Dilemma: The Budget Approach. Special Report*.

--- (2009c) *Factsheet 3/2009 The WBGU Budget Approach*.