

**THE MEASUREMENT OF ECONOMIC PERFORMANCE
AND SOCIAL PROGRESS REVISITED**

OFCE

N° 2009-33

DECEMBER 2009

Professor Joseph E. Stiglitz

Chair, Columbia University

Professor Amartya Sen

Chair Adviser, Harvard University

Professor Jean-Paul Fitoussi

Coordinator of the Commission, IEP

The Measurement of Economic Performance and Social Progress Revisited

Reflections and Overview

Joseph STIGLITZ, Amartya SEN and Jean-Paul FITOUSSI

Prolegomena

THE MULTIPLE PURPOSES OF MEASUREMENT

1. *In February 2008, the President of the French Republic, Nicholas Sarkozy, unsatisfied with the present state of statistical information about the economy and the society, asked, Joseph Stiglitz (President of the Commission), Amartya Sen (Advisor) and Jean Paul Fitoussi (Coordinator) to create a Commission, subsequently called “The Commission on the Measurement of Economic Performance and Social Progress” (CMEPSP). The Commission’s aim has been to identify the limits of GDP as an indicator of economic performance and social progress, including the problems with its measurement, to consider what additional information might be required for the production of more relevant indicators of social progress, to assess the feasibility of alternative measurement tools, and to discuss how to present the statistical information in an appropriate way.*
2. *The Commission reviewed the purposes to be served by systems of measurement of economic performance and social progress, and the extent to which current systems serve these purposes.*
3. *The purposes of our statistical systems are multiple, and a metric that is adapted to one purpose may be ill suited to another. Sometimes confusion is engendered when a measure adapted to one purpose is used to highlight another. For example, GDP is neither a measure of income nor a measure of well-being. What we want to measure is the key question. We may want to measure, for instance, the levels of market activity—one of the original objectives of national income measurement. But increasingly, there is a demand to go beyond measures of market activity to measures of well-being. Whatever it is that we want to measure, we can measure the flows (for instance, the level of production and/or income) during a given time interval, say a year. We may measure changes in those flows, say their rate of growth over time. We may want to compare levels of income across countries or we may be interested in international comparisons of countries’ growth rates.*
4. *It is important to distinguish between these different purposes, since they are affected differently by the various shortcomings of national accounts. For example, measuring the level of production requires a direct measure of government output. But while the adoption of direct measures of government output has been used to give a better measure of growth rates, to date there have been few attempts to adjust the measured levels of national income. To measure changes in levels through time requires distinguishing price and quantity effects, which in turn implies the capacity to measure quality changes.*
5. *There are long recognized problems in GDP as a measure of economic performance, but many of the changes in the structure of our society have made these deficiencies of*

greater consequence. At the same time, advances both in our conceptual understanding of the issues and data availability mean that it is now possible to construct better indicators. Better indicators might be able to address one of the concerns that motivated this report: a marked distance between standard measures of important socio economic variables like economic growth, inflation, unemployment, etc. and widespread perceptions

6. *International comparisons of levels and more importantly of rates of growth play a very important role in the design of policy. Comparisons are indeed possible if the procedures and definitions used to compute the accounts are comparable. Yet there are still “large differences in the ways National Accounts calculations are carried out even among European countries, let alone between Europe and the U.S¹”. This may have far-reaching consequences. It makes no sense, for instance, to structural reforms intended to import the “best practice” of the country performing the best in terms of growth rate, if the growth rates of the two countries differ mainly because of differences in the ways National Accounts are computed.*
7. *Policies have also been affected by widespread statistical analyses of the determinants of growth and economic performance; but inferences made from those statistical studies may be flawed if the measures themselves are flawed. Those conducting this research and relying on these results must be well informed concerning the limitations of our statistics on which they are based.*
8. *But what is of particular concern is when narrow measures of market performance are confused with broader measures of welfare. What we measure affects what we do; and if our measurements are flawed, decisions may be distorted. Policies should be aimed at increasing societal welfare, not GDP. Choices between promoting GDP and protecting the environment may be false choices, once environmental degradation is appropriately included in our measurement of economic performance. This report, building on extensive earlier work, describes the additions and subtractions that can and should be made to provide a better measure of welfare.*
9. *Advances in research across a number of disciplines enables, however, the development of broader, more encompassing measures of well being. Some of these dimensions are reflected in traditional statistics, but are given more prominence: unemployment has an effect on well-being that goes well beyond the loss of income to which it gives rise. Other dimensions to which we call attention are health, education, security, and social-connectedness. These dimensions affect the capabilities of people which depend on the extent of their opportunity set and of their freedom to choose among this set, the live they value². Some economic reforms in recent years may have increased GDP, but their adverse effects on these other dimensions on Quality of Life are unmistakable.*
10. *We care about the future—that the living standards that we enjoy today should be enjoyed by future generations. Our statistical systems should tell us whether or not what we are doing is sustainable, economically, environmentally, politically, or socially. There is*

1. Joachen Hartwig (2005): “On Misusing National Account Data for Governance Purposes », Working paper 05-101, KOF Swiss Economic Institute, ETH, Zurich.

2. Amartya Sen: “Well Being, Agency and Freedom: the Dewey Lectures, 1984”, *The Journal of Philosophy*, 82, 169-221, 1985.

reason to believe that, at least in certain dimensions, what we are doing is not sustainable, but current statistics do not reflect this—just as they gave little indication of the unsustainability of the U.S. economic growth in the years preceding the crisis.

11. *There is no single indicator that can capture something as complex as our society. But because what we choose to measure and how we construct our measures can have such an important role in the decisions that are made, it is important that there be an open and public discussion of our system of metrics. Hopefully, this Report will play a role in this Public Dialogue.*

I - Uses and Misuses of GDP

1. Between the time we began working on this Report and the time we completed it, the economic context has radically changed. We are now living through one of the worst financial, economic and social crises in history. Part of the reason why the crisis took many with such surprise is that the “measurement” systems we use to assess and monitor economic performance failed. They suggested that in the years prior to the crisis the economy was doing far better than it was. The crisis has raised questions of how to value assets – if we value the houses produced back then at current market prices, output at that time would be much lower. So too for the profits recorded by one of the seemingly most dynamic sectors of the economy, the financial sector, which were not just ephemeral, but, in many cases, based on suspect valuations. Not only have questions been raised about the assessment of the economic performance of these firms, but because GDP, our standard measure of national economic performance, reflects the sum total of the economic performance of all firms in the economy, questions are being raised about the assessment of national economic performance.
2. No single measure, or even a limited set of measures, can provide all the information required to assess and manage an economy. But many are asking today, why did neither the private accounting system nor the public one deliver an adequate early warning? It is perhaps going too far to hope that if we had possessed a better measurement system, one that would have signaled problems ahead, government might have taken early measures to avoid or at least to mitigate the present turmoil. But perhaps had there been more awareness of the limitations of standard metrics, like GDP, there would have been less euphoria over economic performance in the years prior to the crisis; metrics that incorporated assessments of sustainability (e.g. increasing indebtedness) would have provided a more cautious view of economic performance. In effect, sustainability issues, in a broad sense, apply not only to the environment and to natural resources, but also to other types of capital.
3. Often, the requisite information is available somewhere within our statistical apparatus, but because it is not given sufficient prominence, it is overlooked. To some extent, that was true for this crisis: data about increasing household indebtedness was available³.

3. Although, the level of housing prices was so high, that the inference that this indebtedness was sustainable was common.

But the excessive focus on GDP metrics—and a failure to understand the limitations of that measure-- meant that these statistics were not given the prominence they deserved. At other times, information is subject to alternative interpretations.

4. Even prior to the crisis, many concerned about the degradation of the environment or the depletion of natural resources argued that national accounts gave a false impression of economic performance, as they measured the “goods” being produced, but not the “bads”. They worried that growth statistics did not include an adequate appraisal of its sustainability⁴. If we take a broader view of sustainability and include other assets, such as human capital and social cohesion, information is simply missing.
5. This dramatic episode is teaching us a very important lesson: Those attempting to guide the economy are like pilots steering a course without a reliable compass. The decisions we make depend on what we measure, how we do our measurements, and how we interpret them. We are almost “flying blind” when the metrics on which action is based are ill-designed. Today, there is a broad consensus that we need better metrics and that we need to understand the limitations and uses of existing metrics.
6. Statisticians, economists and governments have, of course, long been concerned with the many shortcomings of our statistical systems. Many of the issues we will discuss in this overview are longstanding issues, which were already raised by those very scholars who helped to build the present System of National Accounts. The joke of Professor Paul Samuelson (one of the great modern economists, and winner of the 1970 Nobel Prize) is known by almost all generations of students since the 1950s: “what happens to the GDP when a professor marries his servant?” William Nordhaus and James Tobin of Yale University (winner of the 1981 Nobel Prize) in a famous paper published in 1973⁵ highlighted many of the problems that still lie before us. Since then, these problems have been further investigated, and we have drawn extensively on the large and important literature that deals with them.
7. Moreover, reflecting many of the same concerns which motivated this Commission, work has been undertaken on the measurement of economic performance and social progress by the OECD, EU and UN. The references to the literature and to these works appear in the technical chapters of this Report, and they have been an important source for the work of the Commission. It is not just that the Commission’s work does not exist in a vacuum; it is precisely because there is such interest in these issues that we believe that the timing of the work of the Commission is opportune. We hope our Report will provide further impetus to meet the difficult tasks ahead.
8. In this overview we will attempt to give a flavour of these problems, while at the same time emphasizing the difficulty of constructing alternative measures. In writing our Report, we do not underestimate the sophistication and seriousness of current national income accounting practices, nor the difficulties facing attempts to make improvements. There are nearly always good reasons for the imperfect compromises that are made in constructing the accounts. But while those compromises may make sense for one

4. A very small number of countries have wealth accounts. The third chapter of this Report tackles the issue of valuing assets (wealth) and changes in wealth, which is at the core of assessing sustainability.

5. Nordhaus, W. and J. Tobin (1973), “Is Growth Obsolete?” in *The Measurement of Economic and Social Performance*, National Bureau of Economic Research, 1973

purpose, they may not be the best or most appropriate compromise for others, as the following discussion should help to clarify. And those who use these statistics, whether they are scholars attempting to understand better the determinants of economic performance or politicians attempting to shape policies that will ameliorate existing societal problems and lead, more broadly, to social progress, must understand better the limitations of these indicators.

9. The purpose of this Report is to take stock of what has already been done, to assess the weaknesses (and strengths) of current metrics, and to propose some modifications to the existing statistical apparatus. In doing so, we attempt to reflect the growing concerns of citizens throughout the world about quality of life – its objective as well as its subjective determinants⁶ – and the sustainability of growth and the environment. We hope too that this Report will promote greater understanding of the assumptions that underlie the construction of these metrics as well as their consequent limitations, and that it will give a renewed impetus for the many initiatives attempting to improve our metrics and the data on which they rest.
10. Statistical systems are tools for economic management, but they serve a multiplicity of objectives. Our statistical systems were originally created to provide an assessment of the performance of the *market economy*, and, particularly, to monitor cyclical fluctuations. This is still a critical function. But the metrics developed for this purpose are often used (or misused) as a measure of societal well-being. This would be justified only under a set of very specific – and most often unrealistic – assumptions. One of the purposes of this Report is to bring to the fore the problems in doing this, and to suggest alternative measures that might more adequately reflect broader notions of societal well-being and social progress. National income accountants have already, in fact, constructed a variety of measures, and we hope our Report will provide some guidance to the appropriate use of each, along with greater awareness of their respective limitations.
11. In our review of existing metrics, we will discuss (a) some of the areas in which our metrics are incomplete, leaving outside measurement some important phenomena like the environment, home production etc.; (b) the defects of some existing measures. While sometimes the information required to construct a better metric is not available, in other instances it is.
12. Thus, our Report is simultaneously a plea for revisions of our national account system, a call for the end of the dominant use of GDP in assessing everything – performance, well-being, quality of life, etc. – and an appeal for the gathering of more information that would enable us to assess and monitor economic performance and social progress better, so as to reflect those things that citizens care about.

6. The development of the capability approach by Amartya Sen emphasizes the objective determinants of the quality of life, without reducing the measure of the quality of life to those determinants. Cf. Sen A. (1999), *Development as Freedom*, Oxford University Press, Delhi.

I-1 Measuring Economic Performance

13. It has become a commonplace to say that it is important to monitor and assess performance, whether of firms or countries or individuals. This is especially the case as our societies have become more performance-oriented. We expect results, whether from our managers, our workers, or our politicians. Individual rewards are typically based on performance, and incentive systems have to be based on metrics. What we measure affects, of course, what we do. And what individually or collectively we are aiming at affects what we measure. There is an intricate relationship between objectives, measures and actions. If teachers are rewarded for their students' performance on reading test scores, they will teach reading, perhaps at the expense of broader cognitive skills.
14. So too for countries. Politicians are asked to guide the collective actions of society, whether through regulations, taxes, or expenditures. But towards what ends? They are told to increase GDP. But they are also told to pay attention to many aspects of the quality of life – from social justice and urban amenities to noise, air, and water pollution. The two sometimes seem contradictory – paying attention to social objectives sometimes seems to run counter to pursuing economic objectives. But that should not necessarily be the case. Economic activities are not so much an end in themselves, but a means to an end – to higher living standards. If our indicators suggest that pursuing actions directed at improving living standards, broadly defined, have an adverse effect on the economy, perhaps the problem is with our economic measurements.
15. The apparent contradiction between the different objectives alluded to in the preceding paragraph is as old as the post-World War II system of National Accounts itself. In effect, while national income measurement had its origin in macroeconomic accounting, stemming from the work of Keynes and Kuznets, there was another, often more theoretical literature, which focused on the measurement of welfare, and whose theoretical underpinning may be found in the work of Pigou, Hicks and Samuelson.⁷ There is tension between these two strands, which may lead to conflicting views about the treatment of some aspects of the economy, for example, the government sector.⁸ But this tension has perhaps increased through time as the economy became more complex, and measurement of non-market phenomena – in particular, the increased provision by the state of public goods and services – became more important for a fair account of economic activity. But if the public debate and a number of remarks in this overview are leaning towards welfare measurement as a primary purpose, that does not mean that we propose to dispense with macroeconomic accounting. On the contrary, we think that the role of accounts in macroeconomic management is, has been, and will remain essential. Indeed, we have entitled the first part of this Report, in which we discuss how we might improve the measurement of the economy's output "Classical GDP issues".
16. There are other reasons for dissatisfaction with our system of economic measurements. They often seem out of touch with our perceptions. Government statisticians may

7. Arthur C. Pigou (1920), *The Economics of Welfare*, MacMillan; John R. Hicks (1940), "The valuation of social income", *Economica*, vol.7, pp. 105-24; Paul A. Samuelson (1947), *Foundations of Economic Analysis*, Harvard University Press.

8. There is a deeper conflict – most of the literature in the Pigou tradition assumes markets that function well, while the macroeconomic literature is concerned about valuation in situations where there is a deep "market failure", which can lead to persistent unemployment and excess capacity. These market failures play a central role in the discussion below.

correctly announce that the rate of inflation has been “only” 3%, but an important fraction of citizens may feel that their own costs of living are going up far faster. They may be told that the country is experiencing a rapid increase in economic performance, and yet there may be a widespread perception that living standards are going down.

17. Such disparities between what government agencies say and what individuals feel or know about their own status contributes to public distrust of government, and of societal institutions more generally. If the measurement system which is supposed to give a broad account of the situation is incongruent with people’s own perceptions, they may have the feeling that the statistics that used to describe what is happening in our society are being manipulated, in an attempt to manipulate the democratic processes, and as a consequence they may lose confidence in democracy itself. (We should emphasize that, in most cases, the government is not in fact engaged in such manipulation; it is simply that, especially in circumstances where there is less than full confidence in government, views that this is the case may become widespread.⁹) That could have far-reaching consequences. It is why governments in many countries and especially in Europe have taken the problem seriously, and rightly so. In the last few years, a number of official reports have addressed the measurement of inflation and living standards to respond to the heated debate about their measurement and about alleged underestimations by statistical offices.¹⁰ Other countries have attempted to construct broader measures of social progress that are more reflective of societal values, paying more attention to the environment, for instance.¹¹
18. It is thus important to know the reason for these disparities, and to address them. There may be a simple explanation – the benefits of an increase in GDP may be going to a relatively few individuals, with most individuals in society actually being worse-off. If that is the case (and it is, in many countries¹²), then it suggests that greater attention be paid to metrics of societal performance other than average per-capita income, to metrics that are more congruent with widespread perceptions of well-being. It is theoretically easy to construct such measures. Median disposable income describes the income of the “representative individual” – representative in the sense that half the individuals have a higher income, half a lower income. It is thus a good candidate. But as Chapter 1 of the main Report shows, such a construction is complex as it has to rely on data other than those used in the national accounts, and it may be difficult to make this data consistent with the underlying national accounts. Nonetheless, it is essential to create such

9. Of course, there may have been attempts to manipulate accounts. There are allegations that in Argentina under government pressure, the official statistical agency published inflation rates that were much lower than they actually were. Today, the government has to address a very serious credibility problem in the release of economic performance indicators. In other cases, decisions to reform national accounts are in part motivated by the consequences of flawed measurements; overestimating inflation may lead to larger increases in payments that are indexed to inflation. When governments seek to reduce budgetary outlays because of large budgetary deficits, it is understandable that those adversely affected by the statistical reforms become suspect, whether or not those suspicions are justified.

10. Controversies over the measurement of inflation have been particularly intense in the United States (see, e.g. the so called Boskin Report: Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996), “Toward a More Accurate Measure of the Cost of Living”, Final Report to the US Senate Finance Committee, December 4.) and in Argentina.

11. Particularly noteworthy is the work of Bhutan’s government, which has been active in its search for better measurements of societal well-being, as it has attempted to develop a measure of GNH, gross national happiness.

12. A recent study by the IMF made this point very clearly: “Based on observed movements in Gini coefficients (the most widely used summary measure of inequality) inequality has risen in all but the low income country aggregate over the past two decades, although there are significant regional and country differences”. *World Economic Outlook*, October 2007.

measures, if we are to obtain a picture of what is happening to most individuals in our society. (More generally, there are many reasons why it is desirable to try to construct income distribution data. This has been done recently by a few statistical offices, in particular France's INSEE.)

19. On the other hand, the explanation for disparities in perception may be more complex. There may be technical problems with how we measure health and education, two sectors of increasing importance (together, in the US, they comprise close to a third of GDP); or it may be that our measures leave out something important, like our sense of security, or include some expenditures that individuals do not think of as important to their sense of well-being.¹³
20. The risk is that, as countries strive to increase *measured* GDP, they take actions which now, or in the future, may actually lower societal well-being. This is especially the case if our metrics do not take account of *sustainability*, if current consumption puts in jeopardy, for instance, future living standards. The most obvious cases involve depletion of resources and the degradation of the environment. Countries that enjoy high living standards today by depleting their inheritance of natural resources – without investing the proceeds – are “robbing” future generations. It is possible that doing this does not even increase their welfare, as people usually care about the well-being of their children, but they may unintentionally act this way, at least partially because they are not informed, absent the right metric.
21. The world as a whole has put the planet at risk, as emissions of greenhouse gases lead to global warming, with potential catastrophic effects, at least for some countries. We know that the way we, as an international community, have conducted ourselves is not sustainable. We know that it is impossible for the world, as we know it, to survive if current patterns of living and production continue, and even more so if they are extended to the billions in the developing world. Yet, some, even in the richest country of the world, claim that were they to change, we would pay an economic price. Clearly, our metrics are faulty: our measures should tell us that what we are doing today is not sustainable, that current consumption is at least partly at the expense of future generations; in that sense we may be living beyond our means. Better metrics would indicate higher *sustainable* incomes from altering patterns of consumption and production in ways that reduce emissions.
22. This is not the only example of non-sustainability. Argentina's overall growth in the early 1990s was based, in part, on the growth of consumption, and the growth of consumption was based, in turn, on borrowing from abroad. On the basis of standard metrics, it *appeared* as if the economy was performing well. Yet better metrics would have reflected the increased indebtedness; it would have suggested that the country's growth might not have been sustainable. So too the US, and indeed many of the advanced industrial countries, experienced unsustainable growth in the middle of this decade. Again, growth was based on borrowing. In the US, the borrowing supported an unsustainable housing bubble that led the country to believe that it was wealthier than it

13. There may be “psychological” problems: in forming their beliefs about inflation, individuals may pay more attention to increases in the products of certain commodities that seem particularly salient than their weight in their market basket might indicate. Later, we will emphasize that individual's sense of well-being may be affected by being unemployed in a way that is greater than can be accounted for by the loss in income.

was and that it could accordingly live beyond its means. It borrowed hundreds of billions of dollars from abroad, not for productive investment, but for a consumption binge. The standard metrics, however, provided no warning sign, no reliable measure of sustainable consumption. Not surprisingly, nothing was done to curb the excesses. One may object that these episodes were the consequence, not of the metric by itself, but of the fact that in both cases the relevant metric was ignored. After all, data on indebtedness are not lacking! The analysis was faulty, not the metric. But in fact in both cases the data were also telling us that the wealth of the country was increasing and thus that growth was sustainable. This emphasizes the need for better measures, or indicators, of sustainability.

23. One of the consequences of using incomplete metrics may be wrong inferences about what are good policies and practices. While greater reliance on quantitative methods in economics and other social sciences has led to increasing use of metrics in analytic studies (e.g. to assess the determinants of good economic performance), not all those who employ these data are fully aware of the assumptions that go into their construction, the complex trade-offs that inform these decisions, and their consequent limitations, or of the implications that these limitations might have for the inferences that are drawn. Users of these studies are in turn often even less aware of the implications of these limitations for making inferences about alternative policy regimes.
24. For instance, throughout the world, many economists are engaged in studying what makes for successful growth. There are large numbers of empirical studies relating economic outcomes to various policies. If the wrong metrics¹⁴ are employed, incorrect inferences will be made. Many concluded, for instance, that financial deregulation was good, because it led to rapid expansion of the financial industry and an increase in *measured* GDP. We now know that that growth was not sustainable; that much of the profits earned in 2004-2007 might more appropriately be looked at as winnings in gambling by some, which were more than offset by the losses in 2008, and the following years, by others. When looked at from a five-year perspective, the sector earned zero or negative profits. Better metrics will not ensure that we always devise the *right* policies or even that we make the right inferences, but they do make it less likely that we will make the *wrong* inferences.
25. The fact that the financial sector was growing so fast should have been a warning sign. Financial services are, for an important part, a means to an end, not an end in themselves. They are supposed to improve the economy's ability to manage risk and allocate capital, and in doing so, they increase its overall efficiency, capturing for themselves a fraction of the gains. To a large extent, the benefits of the financial sector should be reflected in the increased output of the economy, especially in the long run, as better management of risk enables the economy to undertake higher risk while obtaining a higher return on activities, and as the improved allocation of capital leads to increased growth. The fact that a sector which largely provides intermediate goods was growing

14. Here as elsewhere in this overview we will encounter the same problem: it is not easy to disentangle what is due to the imperfections in existing metrics and what is due to faulty analytical interpretations and/or misuses of existing metrics. This problem is not surprising as we need a conceptual framework to build a measurement system. We hope that our Report will contribute both to the construction of better metrics, which will lead to more reliable inferences, to greater attention to the use of the appropriate metrics (within the set of existing and/or improved metrics), and to greater awareness of the limitations of the metrics, and therefore of the inferences that can be drawn from studies based on their use.

so rapidly should be a cause of concern, unless our metrics show a more than compensating improvement in the measures of economic performance that focus on final output – on the goods and services that individuals actually enjoy.¹⁵ If intermediate goods and services are taking up more and more of a society’s resources, it could be a sign of increasing inefficiency (except if the country is selling those services to other countries).

26. In several countries, the issue of the measurement of *economic performance and social progress* has become especially important, precisely because of the worry that standard measures may encourage our societies to move in wrong directions. And as the present crisis is telling us, moving in wrong directions can lead to social distress and a deterioration in social welfare.
27. Developing countries may be encouraged to allow a foreign mining company to develop a mine, even though the country receives low royalties, even though the environment may be degraded, and even though miners may be exposed to health hazards, because by doing so GDP will be increased. But this Report emphasizes that GDP is not the only measure on which the country should be focusing. It should focus also on the sustainable well-being of the citizens of the country, and even though GDP may be increased, better and/or alternative measures may show the mine as decreasing societal well-being.
28. Today, as we have noted, many countries face difficult choices as they begin to address a range of environmental concerns. This will require changing not only modes of production, but also patterns of living. There are high environmental costs associated with suburban sprawl. There will have to be heavy investments in better-insulated housing and more fuel-efficient cars. Some will complain that the economic costs are too high and that as a consequence measured GDP may go down – but at least some of those complaints are based on a measurement failure: our metrics do not accurately reflect economic well-being, nor the fact that such expenditures may actually increase well-being in the long run.
29. This Report attempts to provide a more systematic account of the deficiencies in our accounting systems and to suggest a research program that might address them. Most of the examples given below have long been recognized by specialists in national income accounting, but we thought it nonetheless useful to review them, partly because they may not be well understood by many of those who make use of these metrics and/or of studies based on them. Some of the anomalies in our current system are small and/or technical, but others go to the heart of what we mean by “societal well-being”.

15. Some of the output of the financial services sector may be a final good in conventional accounting, e.g. that part that is sold directly to individuals (rather than to firms.) Even then, of course, in measuring improvements in economic performance we should be measuring, for instance, the savings in overall transactions costs (time and money). Several governments have argued that revenues generated through certain fees may also reflect monopoly power; as we explain below, there are difficulties in ascertaining how best to include such fees in our system of national accounts. Further problems may result when there are persistent problems of consumer ignorance or producer misrepresentation. Assume that individuals are not fully apprised of the fees that they pay on their credit cards. If they fully understood them, they would not have borrowed as much on their credit card. The “exploitive” fees do not then really represent their valuation of the services provided by the credit card company. Under current account, these fees are an addition to GDP. They might more accurately thought of as simply a transfer payment from the household to the finance company. An increase in these exploitive fees does not represent an improvement in economic performance; it is the contrary.

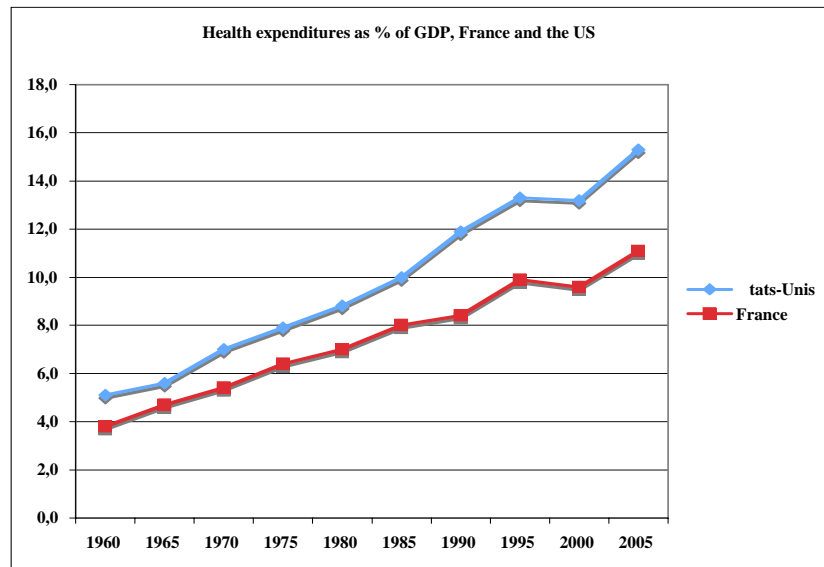
I-2 On some complexities in the measurement of economic performances

30. There is a long list of ways in which standard measures fail to capture important aspects of societal well-being. A standard joke captures some of what is wrong with our measures. Contrast the impact on GDP of two individuals. One is a happily married woman (man) who goes home after work to her husband (his wife). Both may go home after a working day. They get pleasure from cooking their gourmet meal together, using ingredients which they have grown in their garden, and which they follow by a quiet evening reading together. The net contribution to GDP is the value of the few ingredients in their meal that they had to purchase and the cost of the books. By contrast, the lonely bachelor eats an unhealthy meal at a fast food restaurant, then goes to a bar where he drinks excessively as solace for his loneliness, visits a prostitute, and then wrecks his car while driving back home, taking a taxi the remaining distance. This unhappy individual has, by contrast, contributed greatly to GDP – the cost of preparation and serving of the meal and drinks, the sexual services, the repair costs of the automobile, and the taxi home all enter into GDP accounting¹⁶.
31. An increase in violence in society decreases our sense of security. If we respond by building more prisons, by hiring more bodyguards, by installing more surveillance systems, GDP may go up. But no one would say society is better-off. Those who are spending more to protect themselves against increased violence are suffering a loss of welfare both because their purchasing power on other goods decreases and because they feel insecure. This is an issue of considerable importance, as expenditures on prisons (and security costs in general) have been rising rapidly.¹⁷ Who would pretend that growth in the manufacturing and trade of private guns, mainly for self-defense reasons, is contributing positively to GDP?¹⁸
32. The US spends more on health care (as a percentage of its income and per capita) than any other country. In 2008, health care expenditures represented more than 15% of US GDP, compared to 11% in France. While the percentage of national income spent on health has been increasing around the world, the gap between the US and France (and most other countries) has been increasing, as shown in the following figure.

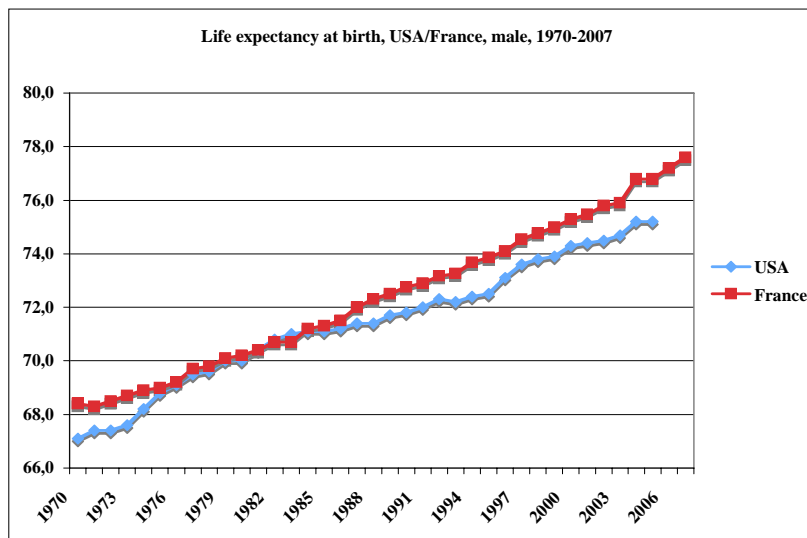
16. Of course this joke is willingly caricatural as its purpose is pedagogical. Sexual services are usually part of the underground economy, as they are illegal. Homework falls disproportionately on women; and some of the more productive home activities, like childcare, are not considered in the joke.

17. Between 1986 and 2001, State prison expenditures in constant dollars increased at an annual rate of 6.4% compared to 4.2% for education. It is the highest rate of growth of all state expenditures. Cf. "Bureau of Justice Statistics, Special Report on State Prison Expenditures, 2001", June 2004, NCI 202949. <http://www.ojp.usdoj.gov/bjs/pub/pdf/spe01.pdf>.

18. This example illustrates a problem that we will discuss more extensively later: some might claim that the increased purchase of guns has increased security from what it would have been otherwise, and in that sense, they are contributing to societal well-being. From this perspective, it would be wrong to exclude them from our measurement of national output. But if the reason for the increase in the purchase of guns is that there has been a general increase in violence, then the change in societal well-being would have to take that into account. The increase in arms purchases will only partially offset the effect of the increased insecurity. In that case, it might seem better not to include these purchases, at least for purposes of making comparisons over time. Indeed, even excluding these purchases results in an overestimate of the improvement of societal well-being, because it does not include any measure of the residual increase in insecurity.

Figure 1- Health Expenditures as % of GDP: France and the US

33. Yet health outcomes are lower in the United States, whether measured by infant mortality, life expectancy at birth, or morbidity. The following figure compares for males the evolution of life expectancy at birth between the two countries.

Figure 2-Life expectancy at birth (male): France and the US 1970-2007

34. We are supposed to be looking for performance (output) measures, and yet traditionally, in standard national income accounts, we use as our performance measure a measure of input (amount spent providing the service). If the US health care system becomes less efficient (as some believe is the case), then inputs go up, and *GDP increases*. Our performance measure should be penalizing the US, and yet it seems to be rewarding it. It is true that constructing output measures is not that easy, but it is not an excuse for overlooking the problem. EU Member States were required to adopt direct output measures under the European Commission decision of 2002. There has been extensive

analysis of productivity changes in the health care sector a number of countries. In the UK, the Office for National Statistics has been investigating measures of health gain to give a quality adjustment to health care output. But the results of all these attempts are still in their infancy, and much work is required if we want to be confident in making international comparisons.

35. The same problem arises in making comparisons of GDP across countries, say comparing US and French GDP. Part of the explanation of America's higher GDP per capita is that it spends more on health. If the US and France had the same health outcomes, then one might argue that health sector "activity" *appropriately measured* should be the same. If we make that adjustment alone, the difference between US and France's per capita GDP is reduced by about one-third.
36. As always, matters are more complicated. Perhaps Americans are naturally less healthy, and therefore it takes more "market activity" to produce the same health outcomes. Our statistics should reflect the "value added" of the economy. If Americans were naturally less healthy than the French, and yet because of more spending on health, Americans succeed in achieving comparable health outcomes, then we should credit the health care system with this achievement: this is its "value added". For example, life expectancy at high ages in the US is one of the highest in the world. But it may not be easy to identify what changes are consequences of the way the economy functions. America is faced with an epidemic of obesity. Some critics blame America's fast food economy and firms that push addictive foods, even on children. Others look towards other societal forces. If one attributes the obesity to economic forces, then additional economic activity in the health care sector directed at correcting the problems created by the distorted market are simply undoing the damage that the malfunctioning economy creates; it would be wrong to include these expenditures as part of our measure of how market activity contributes to an increase in our well-being. But if the origins of growing obesity lie elsewhere, then we should recognize these economic expenditures as part of our measure of economic activity, noting that they are making us better-off than we otherwise would be.
37. An analogy may be helpful. Assume that for some perverse reason an economy hires people to randomly knock down buildings, and then hires people to reconstruct them. A new President gets elected in the country, and he expands economic activity by hiring more people to knock down more buildings, and then hires more people to reconstruct them. In the standard measure, GDP has gone up (twice – because of the additional activity in both destruction and reconstruction). Yet no one would claim that in normal time well-being has increased. (This would not be true in effect if the policy was implemented during a recession and generated more income through multiplier effects.) In this particular example, there is an alternative metric that would provide a far better picture of economic performance: Net Domestic Product (NDP) (rather than Gross Domestic Product) would take into account the destruction of capital. The example shows that the pervasive use of GDP for measuring economic performance may lead us to wrong inferences; in this case, the problem could easily be corrected, by shifting to an alternative metric. But in some of the other instances we discuss in this Report, there is no easy solution.
38. Consider the problems posed by the suburbanization process underway in many countries. As a result of increases in commuting, GDP increases because of greater

spending on transportation. If the country fails to provide adequate public transportation, there will be greater use of private means of transport, and an increase in traffic jams and thus higher oil consumption, all contributing to increased GDP. But the well-being of society is clearly lower. People literally lose time (commuting time is not leisure, nor, like work, does it yield a direct return) and pollution increases.

39. If the road infrastructure is in bad shape, GDP may increase still more, as a result of a further increase in the number of car accidents, with the resulting increase in the turnover of the automobile repair sector and in medical expenditures. What should be treated as a bad – a deficiency in the provision of a public good – will be accounted as a good – an increase in GDP. Again, better accounting systems that take account of the destruction of physical and human capital would avoid these faulty inferences, but it is hard to make the necessary adjustments, and in practice none of the metrics in use do so fully.
40. The worry is that *some* of the increases in economic activity that have been observed in many countries are analogous, though in much less transparent ways.

II. Accounting for the changing structure of our economy

41. There have been large changes in the structure of our societies, which mean that even if the measures which were used 50 years ago did a good job of reflecting economic performance *then*, they may not do as good a job now. These changes in the structure of society – and differences across countries – mean that we need to be careful in making comparisons over long periods of time. By the same token, differences in the structures of different economies make comparisons across countries difficult. In addition, there have been changes in societal values. Even if environmental degradation and resource depletion were as important a half century ago as they were today, there was less recognition of these problems, and therefore less recognition of their import for the measurement of economic performance or societal well-being. While most of the problems to which we call attention in the following paragraphs have long been recognized, these changes in our economy and our society have given them heightened importance.

II-1 The growing role of imputations

42. The step rise in public expenditures after the Second World War and the concomitant expansion of government have in many countries represented a structural change.

Table 1: Total government expenditure as %GDP

	1950	1970	2008
United States	21,4	32,1	38,6
United Kingdom	34,2	42	47,6
France	27,6	37	52,7
Germany	30,4	40,5	44,0

43. But GDP is constructed by adding up *the market value* of goods and services produced in the economy. We add up the apples and oranges and other products produced in the economy using market prices, because market prices are *assumed* to reflect individual's (marginal) valuations of these different commodities. But for a large, and increasing, fraction of the goods and services produced in the economy, the conditions under which such an assumption is true are not likely to be satisfied. This is particularly true for the goods and services produced by government to households on a free basis (mainly health and education). Besides, more and more people live in their own homes. Because they are not paying rent, we don't have a measure of the value of the current housing services.¹⁹ We estimate these "imputed rents" (economists refer to these as imputations; others might call these educated guesses). Table 2 shows that in 2007 these two elements alone – imputed housing, and publicly provided goods and services – constituted 18.4 % of household income in the USA, up from 16.7 % in 1985. In France and Finland, the same imputations are around 30 %.

Table 2: Impact of imputations on household disposable income

% of income	USA		France		Finland	
	1985	2007	1985	2007	1985	2007
Imputed rents	8.8%	10.1%	6.9%	10.1%	9.2%	12.2%
Social transfers in kind	7.9%	8.3%	17.3%	19%	19.5%	22.3%
Total	16.7%	18.4%	24.2%	29.1%	28.7%	34.5%

Source: OECD National Accounts.

44. But the problems do not stop here. If we include housing services provided within the normal household, why not other services that are also not mediated by markets? In Chapter 1 of the main Report, we estimate these services to be between 30% and 40% of conventionally measured GDP. If we include these, then the fraction of our metric which is based on "imputations" rises to about 60%.
45. This illustrates a central problem in constructing metrics: the more comprehensive we attempt to make the metric, the less meaningful may it become, in the sense that more of the numbers are based on imputations.
46. It means, of course, that we need to think very carefully about the *purposes* for which we are using the metric. One of the reasons that GDP measures were originally devised,

19. Somebody living in his own home might enjoy the same housing services year after year, but the market value of those housing services might vary a great deal, reflecting changes in rental markets. Still, in spite of this, there is no change in the individual's standard of living. Well designed statistical systems should reflect this.

as we noted above (par. 10), was to measure the level of *market economic activity*, the goods and services produced *in the market*. This is a rather narrow objective, and even here, there are problems, as we have noted.

47. There are three sets of often interrelated problems: (a) conceptual problems of what should be included; (b) measurement problems of what is included; and (c) valuation problems of what is included. Sometimes we can measure inputs, but not outputs. Sometimes we can observe outputs, but are not sure how to value them. The biases are often complex: sometimes they lead to an underestimate, sometimes to an overestimate.

Intermediate goods

48. In the following paragraphs, we describe some of the controversies over what should be included in our metrics.
49. We explained earlier that we should not include intermediate goods and services (like financial services), but only final goods and services, to avoid double counting. Intermediate goods are in effect already included in the value of final goods and services. But it is not always possible to ascertain what is final and what is intermediate. Transportation services sometimes are part of consumption, but commuting to work is best thought as part of work, a necessary cost to earning income. Typically, we cannot even measure “transportation services”. Rather, we estimate the input of oil and cars used in transport. A more fuel-efficient car will enable the same transportation services to be provided with less input of oil. Unless we correctly account for the improvement in the quality of the car, it might appear as if economic activity is lowered as a result of less input of oil.

Defensive expenditures

50. Individuals are risk averse, and thus value security. Insurance companies provide insurance that reduces insecurity, but many other aspects of economic activity are directed at managing risk. Firms facing the threat of an interruption of supplies may keep greater inventories. Firms facing higher price volatility may lay off some of that risk through trading in futures markets. But changes in economic policy may affect the level of risk, and therefore the need for engaging in these risk-reducing activities. For instance, some argue that financial and capital market liberalization has resulted in greater risk and volatility. If so, it would be misleading to say that the economic activity induced to layoff the risk was an increase in economic well-being. This is an example of a category of expenditures called defensive expenditures, which may in fact constitute an important part of standard measurement of GDP. Here, as elsewhere, the problem is that we don’t have a good way of measuring the output in which we are really interested, individuals’ sense of security. Again, in our metrics of performance (output) we include expenditures directed at reducing insecurity, but we recognize that the way the economy is designed may, in fact, indirectly give rise to the needs that these expenditures are intended to address. (The similarity between this problem and that of obesity, discussed earlier, should be clear.)
51. The problem in making comparisons over time is that the structure of the economy may change in ways that increase (or decrease) the importance of these measurement

problems. Social disintegration may lead to an increased importance of defensive expenditures. Vertical disintegration of production may make the problems posed by intermediate goods more complex to resolve in order to avoid double counting.

II-2 How structural changes may bias our measure of performance

52. A few principles guide our thinking. We wish to avoid obvious distortions and biases. Changes in the economic structure – including in the way certain economic activities are conducted – may result in *systematic* overestimates or underestimates.

a. Biases against home production

53. For instance, our measure of performance should not increase simply because some activity switches from home production to market production or from public production to private production. Recent decades have seen marked shifts from home to market production²⁰; and typically, in developed countries, market production is more important than in developing countries. For instance, people go out to restaurants more often, eating at home less frequently. Even at home, they are more likely to use prepared meals, and when they do not buy a prepared meal, they are less likely to use vegetables grown in their own garden. Few in Europe or America make their own clothes. In traditional extended families, the grandmother babysits; increasingly in the modern family, such services are bought in the market. Fortunately, most national statistical offices now conduct time-use surveys, and we now have relatively good information on inputs of hours into non-market work.²¹
54. This means a measure focusing on market production is likely to overstate increases in well-being: some of the measured increase is simply a *shift* in the locus of production. Indeed, many might argue that no McDonald's meal could come close to being a real substitute for Grandma's cooking.
55. Fortunately, such changes occur gradually over time and for that reason will not cause abrupt changes in measured GDP. If, on the contrary, we see that market output has fallen by 5% between this year and last, it means something significant has happened to the performance of the economy. To be sure, there will be more "leisure" and, of necessity, with budgets cut, more home production of meals. Yet, it says something significant to state that market production has fallen dramatically. ***There are good reasons that measures of market activity play such an important role in our measurement of economic performance.*** In particular, in a recession individual recourse to "non-market" production is not voluntary – it is the result of individuals not being able to find work – so there is typically a large loss of well-being. There is a

20. It might be worth noting that this shift – in the US at least – results from the reallocation of women's time to market work. Men have actually gone the opposite way, devoting more time to home production and reducing their participation in paid employment (although not nearly enough to counterbalance changes in women's time allocation).

21. But while we have measures of time allocation, three problems remain: (a) Some of the time devoted to cooking may be more a form of leisure. (This problem arises, too, in a somewhat different form, in market production; individuals may enjoy market work just as they enjoy non-market work. In principle, to the extent that that is the case systematically in some job, the wage for that job would adjust to reflect these non-pecuniary benefits.) (b) There are difficulties of valuing the time allocated to home production. (See the discussion below). (c) There are difficulties in assessing productivity improvements. We are interested in measures of output; we are attempting to infer output from inputs.

failure of the economy to work in the way that it should. It is important to have diagnostics that tell us of the magnitude of this “disease”, and much of traditional national income accounting is directed at this objective.

56. While it is desirable to go beyond market production, assessing the value of home production poses real difficulties. As difficult as matters are in the public sector (to be discussed shortly), matters are even worse in assessing non-market economic activities. We have difficulty measuring and valuing *inputs*, not just in measuring and valuing outputs. We know that there have been marked changes in home production – automatic washing machines, dishwashers, dryers, and vacuum cleaners are technological revolutions no less important than mechanization within the market. Some of the value of these inputs is “captured” in the value of market sales, particularly when appropriate adjustments are made for the improvement of these household durables. Still, as the discussion below illustrates, significant problems of quantification of these changes remain.
57. Extended families in many countries provide their members “insurance” services. Everybody helps out when someone is in trouble. In some countries, these insurance services are moved into the market, in others they have been shifted to the government. While the “risk” services may be similar, how they are evaluated in the national income accounts may differ markedly.

b. Depreciation and technical progress

58. There have been several other long-run changes in the structure of our economy. Again, taking account of these is not so important when our focus is on the level of market activity (or changes in those levels over a short time-span) rather than on the level of societal well-being. One such change involves depreciation. We need to take account of the depreciation of our capital stock, just as we need to take into account the depletion of our natural resources.²² Otherwise we would overstate the material dimension of our well-being. In a manufacturing economy, depreciation may be largely a matter of machines wearing out. In a modern innovation economy, depreciation is largely a matter of technological obsolescence – which in turn depends on a hard-to-predict variable, the pace of innovation. In the “old” economy, the ratio of NDP (net output, taking account of depreciation) to GDP (gross output, ignoring depreciation) may be high and stable. In the “new” economy, the ratio is probably much lower and certainly much more variable. If this is indeed the case, it means that NDP (and thus a more relevant measure of well-being) has increased more slowly than GDP. More of gross output is spent just replacing equipment that has become obsolete.

c. Imperfections in competition and the New Economy

59. Among the sources of growth in both developed and developing countries have been sectors, like telecommunications and airlines, in which firms can engage in price discrimination. Different individuals may face different prices, and significant quantity

22. These issues are closely linked to those discussed later on sustainability. If we do not invest to offset the depreciation of a machine producing today’s output, that output (and the consumption which is based on it) can’t be sustained. When we add to consumption the gross investment that we are making today, we get an exaggerated view of today’s “output.”

discounts or surcharges may exist. Prices are used to value output, but only because they represent how individuals, *at the margin*, value one good against another. This no longer holds when there is price discrimination and profits no longer measure the (marginal) valuations of the output less the (marginal) valuations of the inputs. If market imperfections grow, then changes in the price of, say, telecommunications services no longer accurately reflect changes in *marginal valuations*. Growth may be overstated.

d. Other market failures

60. Markets fail for a variety of reasons – imperfections of competition discussed in the previous paragraph, externalities such as pollution, incomplete risk markets, or involuntary unemployment. The importance of these market failures can change over time, in some cases markedly so. Urbanization has increased the importance of urban amenities. Waste disposal may not be a problem in sparsely populated rural sectors, but becomes critical in urban areas. We must spend money to get what was once obtained at no cost. In some ways, this is like the switch from non-market activity to market activity; it is a switch from “free” to “purchased”.²³

e. Globalization

61. Globalization is one of the other important forces facing our society, and it too poses challenges to our measurement system. As we have noted, the incomes of a country’s citizens may differ markedly from the output produced within the country – in capital-exporting countries, it may be much larger, in capital-importing countries much smaller. Not taking into account capital flows may provide a vastly distorted picture; and because capital flows can change quickly in a short period of time, this is an arena in which change poses difficulty both for long-run and short-run comparisons. This implies that the relevant measure (at least for the assessment of the welfare of the citizens in the country) in a globalized world is the Gross National Product (GNP), and the refinements to that concept presented in Chapter 1. These measures include net income received from or paid abroad -- rather than the GDP which does not, yet it is the latter concept that is typically used today. For purposes of the welfare of the citizens of a country, what matters is not what is produced within the country, but the incomes of the citizens of the country.²⁴

f. The growth of government

62. Another major change in many economies is the increasing role of government. Again, we would get a distorted view of the world if output were to increase or decrease simply because some activity shifts from private to public production. In principle, it should not change. But in the public sector, we measure output by input. Thus, a shift from the private to the public sector in which nothing else changes almost *necessarily* lowers

23. With some market failures, prices may still reflect individual’s marginal valuations. It is only that the market prices are not those that would prevail were these market failures corrected. Often, however, there are impacts on the well-being of others (this is the case with externalities) that are not captured in market prices; and in many cases there are adverse effects on natural assets, which should be taken into account in any system of national accounts. See below.

24. Obviously, migration, and especially temporary migration, presents further complexities, which we do not address here.

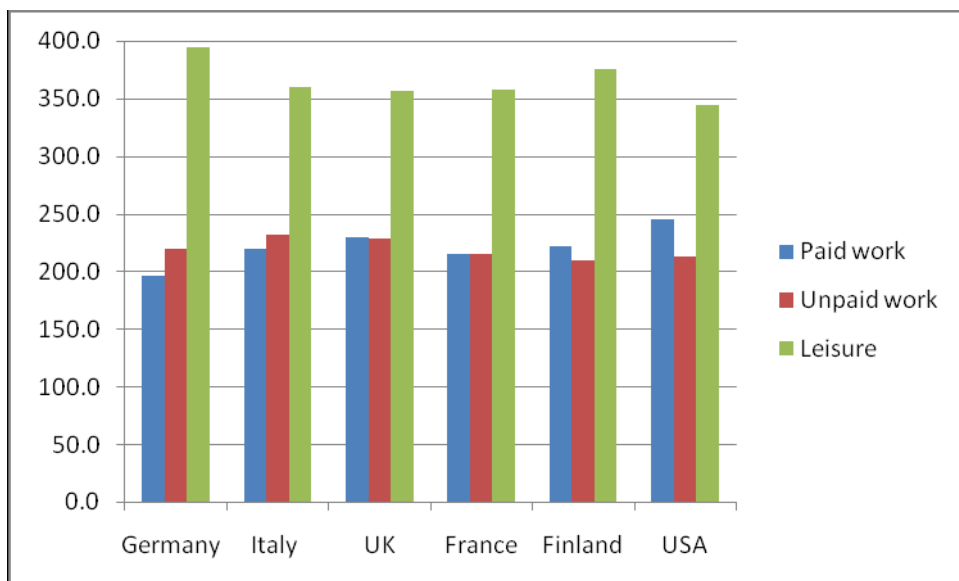
GDP. It lowers it even more if the shift enhances efficiency, because then the input is lowered – and measured GDP falls.

63. While this suggests a statistical bias *against* public provision, matters are more complicated. It is possible that the value of the goods and services provided by the government is less than the value of the inputs. In that case, valuing public goods and services at the cost of inputs overstates the value of what is produced.

g. Leisure

64. Still another change in economic structure involves *leisure*: there are marked changes in the amount of leisure enjoyed over time and amongst countries – and even within countries (a point that we will return to shortly). We have witnessed enormous increases in productivity in the last 150 years. Many economists had expected that most individuals would respond by increasing the amount of both the leisure and the goods that they enjoy. Yet in the United States, on average, Americans have responded in the two last decades by working more and increasing their consumption of market-produced goods and services – in marked contrast to what has happened in many other countries. The table below shows that for the last year available the working time of the American is the highest among the countries considered.

Figure 3- Housework, paid work and leisure
Minutes per day and person, latest year available*



Note. Using normalised series for personal care; Unites States: 2005, Finland 1998, France 1999, Germany 2002, Italy 2003, United Kingdom 2001.

Source: OECD (2008), *Growing Unequal? Income Distribution and Poverty in OECD Countries*, Paris.

65. America's response raises many questions: clearly, our planet could not sustain such consumption levels were they to be extended everywhere, without marked increases in "environmental efficiency", for instance, a reduction in the level of emissions per unit of output. But here again, our focus is narrower: We ask only, what do these differences

imply for evaluating economic performance? Surely, if one society chooses to limit its consumption of material goods, enjoying more leisure, including time devoted to culture, the arts, and community engagement, it should not be counted against it. Citizens in such a society might be far happier than in one which works longer hours, spending less time both with the family and in the community. Citizens in the hard-working society complain that, while they are working hard *for* the family, they have no time left *for* the family. Yet, our conventional measures would attribute better economic performance to the harder- working and unhappier society: both incomes and growth would be higher. Furthermore, the increase in average working time may be itself the consequence of the society's malfunctioning. If inequality becomes pervasive, the number of persons who have to work harder to ensure their living may greatly increase (there is a negatively sloped Hicksian labour supply curve): they may claim that they have no choice but to work harder (though of course they could, were they willing to accept a much lower standard of material consumption than other citizens). It would be questionable whether this evolution is welfare-enhancing, even if GDP increases as a consequence.

66. One needs to look at these issues, of course, from a lifetime perspective, and from this perspective, the discrepancies across societies and over time may be even larger. For instance, in some societies, individuals retire earlier than in others. Mobility may change over time and differ across countries; in countries with lower mobility, for a given level of contemporaneous inequality, lifetime inequalities will be greater. Poorer individuals often have to retire earlier, because of the back-breaking work in which they are engaged; thus their lifetime earnings may be an even smaller fraction of mean lifetime income than their current income is of current mean income.
67. The retirement decision illustrates the complexity of measuring performance. Many individuals would prefer to work more years if their work were interesting²⁵; work is an important part of their self-identity. Standard economic models view work as a *cost*, yet for many individuals, meaningful and decent work gives some sense to life. In some countries, there are important changes to the nature of the workplace. More democratic workplaces can make work more interesting and rewarding. Standard metrics take no account of these changes. The extensions of our systems of metrics discussed in Chapter 2 of the main Report hold the promise of being able to do this.
68. By the same token, individuals spend longer in school today than they used to. Some of this education is “consumption”, but some is “investment”. The former might appropriately be included in “leisure”, the latter as an intermediate good, an input into the production of other goods and services.

h. Distribution

69. A final important change in the structure of many societies around the world is increasing inequality. The distribution of income and wealth determines who enjoys access to the goods and services produced within a society. (It also has import for

25. There is no contradiction here with the preceding paragraphs. Individuals may at the same time value leisure during their professional life and because the latter is rewarding still prefer to retire at an older age.

influence in political processes.) Earlier, we noted that one of the reasons that changes in GDP per capita over time (and differences across countries) often fail to reflect widespread perceptions of such changes (or differences) is that the benefits of growth are not just unequally distributed, but may be becoming more so: most of society can be worse-off, even as a few are so much better-off that average incomes have increased.

70. One of the reasons that most people may perceive themselves as being worse-off even though average GDP is increasing is *because they are indeed worse-off*. This has been the case, for instance, in the United States in recent years: in 2006, median household income (that is, the income of the household such that half of the households had incomes greater than that amount, half less) was lower than in 1998, even though GDP per capita had increased by about 9% over that period. If we take a longer perspective, over the last 40 years median household income has increased by only 30%, while GDP per capita has doubled.
71. While there is no single number that summarizes the manifold distributional changes (or differences across countries), what is clear is that average income does not provide an adequate summary statistic. If societies care about inequality, then a policy that simultaneously increased national income and inequality could either increase or decrease societal welfare. By the same token, if average income is increasing but at the same time inequality is increasing, it is not clear whether societal well-being is increasing or decreasing. And inequality may increase in a variety of ways, as it is a multidimensional concept. The relevant question is: “inequality of what?”²⁶ . If for example, the way chosen to increase production and income, is labour cost reduction through a decrease, say, in health insurance, average income may well go up (because competitiveness is enhanced), but the capabilities of some fraction of the population, their freedom to choose their life will almost surely go down.
72. There are many elements of the distribution of income that constitute part of the picture of what is happening to a society. In societies in which incomes are growing rapidly, elderly people may have a much lower income than younger people, unless societies have a good social security system – beyond that which can be provided by private savings. As we noted earlier, societies with greater social and economic mobility may have less inequality in lifetime incomes than other societies with comparable measured inequalities at a moment of time.
73. What matters is thus not just inequality of income or consumption, but inequalities in the opportunities, capacities, and life chances of those born under different circumstances. The capability approach to the quality of life emphasizes the objective determinants of well-being, without reducing it to those determinants. If the capabilities of different individuals – such as having adequate nourishment or the literacy required to participate actively in political life – become more unequal, social mobility will be impaired²⁷. The question then is why, to start with, capabilities became more unequal and the answer can’t be but that the provision of public goods and services has deteriorated. There is, in effect, considerable concern that in some societies there has

26. Cf. Amartya Sen: “Equality of What?”, in S. McMurrin (ed), *Tanner Lectures on Human Values*, vol I (Cambridge: Cambridge University Press and Salt Lake City, UT: University of Utah Press, 1980)

27. Cf Amartya Sen (op. cit.) and by the same author: *commodities and capabilities*, Oxford: Oxford University Press, 1999

been deterioration in broader measures of socio-economic mobility, though in others it has improved. Policies (such as the quality and inclusivity of public education) may improve these broader measures of socio-economic mobility. This is an arena where the failure to gather relevant data may have led to the pursuit of policies with adverse societal consequences.

74. “In noting the nature of Human lives, we have reason to be interested not only in the various things we succeed in doing, but also in the freedoms that we actually have to choose between different kinds of live”²⁸. This is the philosophical underpinning to the capability approach to measuring the quality of life. It is why this approach, which we describe extensively in Chapter 2 of the main report, argues that the quality of life should be conceived and measured directly in terms of functionings – i.e. beings and doings that people value and have reason to value – and capabilities instead of resources or utility. “Even though the much used economic criteria of advancement, reflected in a mass of readily produced statistics, have tended to focus specifically on the enhancement of inanimate objects of convenience (for example in the GNP or GDP which have been the focus of a myriad of economic studies of progress), that concentration could be ultimately justified – to the extent it could be – only through what these objects do to the human lives they can directly or indirectly influence”²⁹. Focusing on outcome is in effect very limitative, because capability captures not only achievement, but also unchosen alternatives. If a person chooses to work fewer hours to enjoy more leisure, this implies that he had the opportunity to work longer hours. If the cost of unemployment for an unemployed person is much higher than the lost in income, it is likely because it includes a lost alternative.
75. Measurement problems – of the kind addressed in this Report – bedevil our attempt to ascertain what is happening to the distribution of income. We can describe, for instance, how money (or nominal) income for different groups is changing. But that does not tell the whole story. The elderly may consume more medical services, so that if there is a higher rate of inflation in medical services, their real income should be adjusted to reflect this. To achieve the same functioning as the younger, they need a higher income, a characteristic that the capability approach emphasizes. The government may, for that reason, provide more medical services for the elderly without charge or at discounted prices. Our statistics may or may not trace who receives various publicly provided services. We noted earlier that development is often associated with urbanization. But prices of goods and services in the rural area may differ from those in the urban area, and the rate of increase of prices may also differ.³⁰ Recent years have seen prices of housing in urban areas increase markedly relative to rural areas. One of the controversies over the extent of poverty concerns the prices that urban dwellers have to pay for the goods they consume. Locally produced food and housing may be much less expensive, but imported goods may be more expensive. The poor may buy many key goods in small quantities, which drives up their cost of living. Adjustments in “purchasing power parity” numbers in recent years have provided markedly different pictures both of the level of poverty around the world and the magnitude of poverty

28. Amartya Sen: *The idea of Justice*, Allen Lane, 2009, P.18

29. Amartya Sen, Op. Cit.

30. These remarks highlight the risks of using a single consumer price index to adjust nominal incomes for different groups.

reduction. The number of persons in absolute poverty (less than 1 dollar per day) has decreased from about 1.5 billion in 1981 to less than 1 billion in 2004. But if the absolute poverty line is increased to 2 dollars per day, this number has slightly increased during the same period: from 2.4 billion in 1981 to 2.5 billion in 2004³¹.

76. Unfortunately, in many areas, even when we have the appropriate conceptual frameworks, there are problems in getting relevant data. Many inequalities, including those related to gender (both in terms of consumption and household work), occur *within* the household, and therefore are masked by “household” data. There are problems in reconciling data generated by household surveys used in assessing income and consumption inequalities with those underlying national income accounts. Our Report emphasizes the need for greater investments in data.

i. Change in household size

77. There have been many significant demographic changes in recent years that affect the measurement and interpretation of the data on economic performance and societal well-being. Here we emphasize family size. Most individuals live within households, and household sizes have been changing. Most households today are smaller than they previously were. Adjustments are often made to reflect economies of scale in household living. If the average size of households goes down, then, using these household adjustments, well-being will increase less than proportionately to the increase in unadjusted household income. But this conventional approach misses one important point: individuals, for the most part, have chosen to live in smaller households. There are perceived benefits: perhaps fewer “free rider” problems, perhaps fewer coordination problems, or perhaps some people value greatly the increased “space” that it provides. The fact that they could have chosen to live in larger family units suggests that these benefits outweigh the more narrowly defined economic disadvantages.
78. The move to more nuclear families has other consequences. More extended families provide a variety of services (babysitting, cooking) and additional security, which typically are not reflected in our performance measurements.

j. Changes in the pace of change

79. Changes are always occurring, and will always occur. But there have been changes in the pace of change, and these may have a qualitative as well as a quantitative effect. They make the task of assessing change more difficult. They throw into question long-term comparisons. Some have suggested, for instance, that the magnitude of economic fluctuations in the post-Keynesian world is little different than earlier³². But even if the statistical analysis of the *data* is correct, it may be little more than a statistical artifact, a measurement error. We know, for instance, using other metrics, that there is a difference: since the advent of Keynes the length of economic expansions has been longer, of downturns shorter. But with the quickening pace of innovation (if our metrics

31. Cf. Francisco H.G. Ferreira and Martin Ravallion, “Global Poverty and Inequality: A Review of the Evidence”, Policy Research Working paper no. 4623, The World Bank, May 2008.

32. Cf. Christina Romer: “The Prewar Business Cycle Reconsidered; New Estimates on Gross National Product, 1869-1908”, *The Journal of Political Economy*, Vol.97 n°1 February 1989.

on the pace of innovation are correct), there may be problems making comparisons over even shorter periods of time.

Some concluding observations

80. We have listed several important changes in the structure of economies throughout the world. These and other important changes in economic structure pose large numbers of policy challenges. Our concern in this Report is narrower: we are concerned with assessing economic performance and social progress. It may be relatively easy to compare two economies with similar structures. It is far more difficult to make comparisons involving economies with different structures. Since economic structures normally change slowly, these differences in structure may not be that important in assessing the change over a decade. But they take on central importance in assessing changes over longer periods of time, or in making comparisons, even among advanced industrial countries, where, say, one country has chosen to rely more on government provision of certain social services and the other has turned to private provision.
81. The discussion in the following chapters will show the valiant efforts national income statisticians are making to deal with these intractable problems. Yet, it will also show the limitations of these statistics. One of the important lessons to emerge from this Report is the need for caution: one has to be aware of these limitations as one attempts to make judgments, whether about the effectiveness of certain policies, or the magnitudes of progress over time, or differences across countries.

III. The way we proceed

82. Because what we measure affects what we do, the construction of measures of economic performance has become the subject of some political controversy. Not surprisingly, some in the mining industry resisted early efforts to construct measures of “green GDP”, which took into account the effects of mining on resource depletion and environmental degradation.
83. This is a Report of technical experts. But as technical experts, we realize that a society’s values are inevitably embedded in the construction of an index of economic performance and social progress. We have attempted to discuss the construction of indices reflecting the valuations, capabilities, and perceptions of individual members of society. Individuals and households are the basic units of analysis. Yet their perceptions too must be seen within a social context.
84. It is our belief that an open discussion of the issues – and problems – involved in measuring economic performance and social progress provides an important context within which societies can engage in critical debates about societal values. Given the diversity of views within our society, it is not surprising that we have not put forward a single “best” or “correct” way of measuring economic performance and social progress. There are many indicators of societal well-being. But we strive to go beyond simply

listing a set of factors that affect individual and societal well-being. Part of our objective is to find quantitative metrics for key factors, and to assess the relative importance of various factors and to discuss the methodologies by which such quantitative assessments may be made.

85. Consider, for instance, health – clearly an important determinant of individual and societal well-being. There have been marked improvements in health status in most countries around the world. Life expectancy of men in France has increased from 68 years to 77 years over the period 1970 to 2006, and in the US from 67 years to 75 years over the same period.
86. These improvements are an important aspect of societal progress in recent decades. We should make note of these improvements, even if we cannot be sure what causes them, and even if we are not sure of the best way of quantifying them. The changes in complex: there have been large changes in the quality of medicine, and in social mores that affect health status. Changes in the environment or chemicals that are added to food may have mixed effects on health status. But when changes in health status are directly related to economic activity, it would be wrong not to include them somehow in our measurement of economic performance.
87. But how do we add up improvements in health status with changes in material well-being or increases in leisure? Again, there are standard techniques – how much individuals value an additional year of life, or reductions in the risk of certain disabilities. Yet these techniques do not command the assent of many economists and social scientists, let alone support from other segments of society, for whom life may be viewed as priceless. We cannot resolve these matters here; but they suggest that, until there is broader societal consensus, it would be best to present such indicators *separately*.
88. This approach has sometimes been referred to as constructing a “dashboard” of indicators. As is so often the case, there are trade-offs: a single metric is easier to understand, and yet one cannot expect to summarize all the relevant information concerning the performance of something as complex as our society within a single indicator.
89. But while we believe that there should be an array of carefully chosen statistics (illustrated by the discussions of this Report), we believe that the construction of some simple “aggregate” measures, along the lines of traditional GDP measures, can be extremely useful. There is a need for an “extended” or “adjusted” GDP-like measure.
90. This Report progresses through a succession of revisions and extensions of the standard measures. Each of the revisions is important, in its own way, both for assessing changes in well-being and making comparisons across countries.
91. Chapter I of the main Report begins with a review of several long-discussed adjustments. The need for most of the adjustments discussed in Part I is widely recognized, and yet, ironically, some of the reforms of recent years have moved in the opposite direction. In particular, while GDP measures output within a country, much more relevant for assessing the well-being of a country’s citizens is the income of those within the country, which was more closely captured by the GNP numbers that were more widely used prior to 1990. In our report, we discuss refinements to this concept (in

particular, Net Domestic Disposal Income, NDDI). The differences have become particularly important in an era of globalization, with large flows of money going from one country to another.

92. The first set of reforms focuses on *better measures of market income and output* – taking into account, for instance, depletion of natural resources, degradation of the environment (issues which are also central to the discussion of Part III), and better measurement of health, education, and other services provided by the government. The failure to take into account resource depletion means that many countries overestimate their true output – and the failure to take into account environmental degradation almost surely means that the world as a whole is overestimating global income.
93. We proceed from there to look for better measures of individual and societal *welfare*, taking into account most importantly, recent innovations in assessing well-being.
94. The current crisis, and the climatic crisis associated with global warming, has drawn our attention to the importance of *sustainability* – any measure of economic performance (whether it is a measure of output or income or a broader measure of well-being) has to address not just what is happening today, but the consequences of today's actions on the future. The final part of this Report describes the progress that has been made in constructing the appropriate metrics that rise to this challenge.
95. There are accordingly three Chapters to the Report (both in their technical and non-technical versions): The first focuses on improvements in the measurement of output and income (what we refer to as the classical GDP issues). The second on broader measures of well-being; the third on measuring sustainability. Our Report progresses from the easiest to measure to the more difficult, from the areas in which there is broad consensus to those in which there is less consensus. Sections IV through VI deal with these issues. But we are more concerned with well-being than activity. Sections VII will consider other approaches to monitoring societal progress. Finally, in Section VIII we discuss the problems of sustainability.

IV. Measuring the level of market activity and government services

96. We have begun with the challenges posed by measuring the level of *market* activity, proceeding from there to the measurement of overall activity. Sections V and VI consider adjustments to our measurement of market activity that provide a better assessment of well-being (leisure, defensive expenditures, etc.). In a sense, these sections can be thought of as attempts within the classical framework of doing a better job of assessing changes in welfare.
97. One would have thought that there are few problems in measuring market activity. All one has to do is add up the market value of goods and services produced, then construct some measure of inflation to convert nominal values into real values, and one will have a good measure of real national income. As the discussion above should have suggested,

matters are more complicated. Here we highlight a few of the key points that are elaborated at much greater length in the Report.

98. Several of these are related to the fact that, as we have already noted, much of what we call market activity is only partially so. The government buys its inputs (labour, other goods and services) in the market, but most of what it delivers is not sold, and even when it sells certain goods and services, it does not do so in a competitive marketplace. While the discussion below focuses on problems with assessing the value of government output, there are problems in other sectors as well. In many countries, most citizens do not buy housing services, but rather own their own home. While they have purchased their home *in* the market, there is no market price for the housing services they have obtained *this* year.³³

Measuring the value of the output of marketed goods and services: quality

99. Much of the increase in GDP today occurs not as a result of, say, an increase in the *number* of cars purchased, but of improvements in their quality. They pollute the air less, they are safer, they get more miles per gallon, and they may even be more comfortable. The challenge then is to figure out, when the price of a car increases, how much of the increase represents an increase in “quality” and how much inflation. The difficulty is compounded when we deal with services. Whereas for manufactures it is easy to define the unit purchased, this is much harder for many services. This difficulty has long been recognized, and it implies that for the service sector, especially for complex services, the quality question will be much harder to resolve. There are standard techniques – hedonic prices, by which statisticians determine how much individuals are willing to pay for a specific dimension of quality, for instance, an increase in miles per gallon.
100. But this refers to *private*, not social valuations. This highlights one point, made earlier – prices represent *private* valuations, not social valuations.

There is a distinction between private and social valuations whenever there is an externality. If, for example, oil is being subsidized, if there are no charges for greenhouse gas emissions, the private value of a more fuel-efficient car may be low, while the social value may be high. In this accounting framework, it appears that there is a trade-off: tightening the environmental regulations for car producers shows up as a reduction for the consumer in the value of the cars produced, as the car makers have to charge more for something that consumers directly value little. But if we charged everyone for the value of greenhouse gas emissions – if we made users of cars pay for the full value of their “externalities”, their contribution to global warming – then, of course, the *private* valuation of these emission reduction changes would be high. There would be no trade-

33. Systems of national income accounting typically impute the flow of housing services. In principle, if rents fall dramatically (as they are in the United States today), the value of the flow of housing services will fall. From a (short-run) welfare perspective, the effects are offsetting, the lower (imputed) rental income as homeowners being offset by lower rent payments. Neither is there any real change in economic activity, since individuals are doing exactly the same thing as they were doing before. There is a change in the market value of these imputations. If, somehow, the economy managed to get itself through this episode of declining rents without any change in employment and output of other goods, it would be misleading to suggest a marked decline in GDP simply because of the decline in the value of imputed rents.

off. Indeed, if, for some reason, car manufacturers were reluctant to produce low emission cars, then regulations forcing the production of low emission cars would show up as an increase in GDP. *This discussion highlights that many of our perceived trade-offs between environment and growth are little more than a statistical artifact.*

101. Underestimating the extent of improvements in quality is equivalent to overestimating the extent of inflation – highlighting the close link between the concerns of this Report and other reports focusing on the measurement of inflation. For instance, an important, but controversial, report reviewing the measurement of inflation in the United States, the so-called Boskin Commission Report,³⁴ argued that the US had been overestimating the rate of inflation by between 1 and 2%, and thereby underestimating the real rate of growth. Of that amount, it estimated that 0.6% was due to mismeasurement of quality changes. Whatever its limitations, the Boskin Report had the merit of drawing attention to the quality question, attempting to quantify its importance, and proposing measures to deal with it.
102. In the coming years, especially difficult problems will be arising as we start to price certain externalities, most importantly greenhouse gases. It might appear that charging for greenhouse gases raises prices (assuming those charges are passed on) and is therefore inflationary. If firms respond by investing more to produce the same *real output, say of cars*, with less emissions, it will appear as if real output and productivity has declined. Part of this decrease may be explained by the fact that the new cars are more expensive to produce. But the output of greenhouse gases has a social cost – it is an output with negative value. We should have been subtracting the value of this “bad” from GDP. *True* output, valuing both the output of cars and the (negative) output of greenhouse gases, may have actually increased. Again, the failure to assign the correct valuation has given us an incorrect indication of what is happening to societal output.

Measuring the value of output of publicly produced goods and services

103. We have no good way of measuring the value of goods and services produced by the government. We often use a simplification – we measure the output by the value of the inputs. But this assumes that the increase in productivity in the government sector is the same as in the rest of the economy, and that there is no “social dividend”. Consider, for instance, what happens if we nationalized a private insurance company, providing the same insurance policy to the holders. The services of the insurance industry when it is private are measured by the inputs of labour (or other factors of production) plus the profits. But if the government gives the insurance away freely, then, by definition, there are no profits. Citizens pay for the costs, of course, through taxes. It is *as if* the government sold the insurance at the market price and then rebated the profits through a social dividend. Because there are no profits, the output of the nationalized industry will be smaller than the output when it was in the private sector. But this is just a fiction of the way we measure national output. There is a very important caveat here: studies that argue that national output is lower in economies in which the government plays a more

34. Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996), “Toward a More Accurate Measure of the Cost of Living”, *Final Report to the US Senate Finance Committee*, December 4.

important role in the economy *may* simply be reflecting accounting conventions. They are not saying anything about reality. To do that, one would have to look at *real* measures of productivity, for instance, cost per telephone line.

104. The import of this is that if there is faster productivity growth in the public sector than in the private sector, our measure underestimates growth, and conversely if there is slower growth. The problem arises when productivity in the public and private sector do not move in tandem³⁵, so real output does not move in tandem.
105. The difficulty is measuring these productivity increases – and while our analysis shows that how we treat these sectors has large consequences, we cannot resolve whether conventional measures are biased up or down, or whether there are systematic biases in assessing well-being in one country relative to another. The fact that the US has poorer health outcomes while spending more money *suggests* a lower level of efficiency in the provision of health services, and taking this into account may markedly change our perceptions of relative standards of living. And this is true, without even taking into account differences in (health) “security”, differences in risks faced by citizens in America versus other countries.
106. Measuring output by input (and therefore omitting “profits”) might seem to bias the output of the public sector downward. It may, however, be true that in some countries government provides public services that are not valued by their citizens or it provides them in such an inefficient way that valuing output by input may provide an overestimate of the value of the services. There may be, in effect, hidden transfer payments in the public sector, and one should never confuse transfer payments with economic activity.
107. There have been some attempts to measure and value output in the public sector, especially in the UK³⁶, prompted by the type of problems we have just exposed. To understand the complexity of the valuation of output in the public sector, consider education. We can ask, *how many students are educated?* If we spend more resources to educate the same number of students, this would suggest that education productivity has gone down. That may or may not be the case. If most of the extra resources go into administration, it may be a sign of inefficient delivery mechanisms. If, however, additional resources go into smaller class sizes, it may improve the quality of education.³⁷ A good measure would reflect this; there would be an increase in output, even though there has been no change in the number of students educated. The worry is that by focusing on the number of students, we build in the assumption that productivity is declining, even when in the relevant sense, it is increasing. Consider a country where immigration increases. Because of cultural and language problems this may require, to maintain “quality”, a more than proportionate increase in the number of professors and/or a reduction in class size. Should we then speak of a decrease in productivity, knowing

35. This assumes that wage growth moves with productivity, which it typically does. In the short run, however, the two may again not move in tandem.

36. Tony Atkinson, “Measurement of Government Output and Productivity for the National Accounts”, Atkinson Review: Final Report, HMSO, 31 January 2005, Palgrave-MacMillan.

37. Cf. e.g. Sid Glibert, “Quality Education: Does Class Size Matter?” Research File, April 1995, Vol. 1 No. 1; Christopher Jespen and Steve Rivkin, “Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size”, *The Journal of Human Resources*, Vol. 44, No. 1, 2009.

that if these special efforts are not undertaken the overall quality of education would decrease?

108. Alternatively, additional resources may go into providing better education to disabled students. In this case, most students will not see any change in quality, but a few will see marked changes (this is an example where distributional consequences are paramount).
109. We can measure the quality of the education of students – performance on test scores – but that too is an intermediate variable. What we are really interested in is long-term economic productivity (and perhaps even good and well-adjusted citizens). This may be hard to assess. But even then, as we have noted, it doesn't fully inform us of the *value added* by the education system, because other societal changes (some related to the economic system) may affect the quality of the inputs, the performance of the students in the absence of the educational expenditures.
110. In the health sector, measures of input may be an even poorer indicator of output. It is not the number of heart surgeries that is of concern, but how successful they are, and there have been marked improvements.³⁸
111. Fortunately, we do have some indicators of overall success (even if they are imperfect measures of value added): in the area of health, life expectancy, child mortality³⁹, etc., and in the area of education, performance on test scores. (Test scores are only relevant to the extent that they correlate with future productivity or broader measures of well-being, but they are a better indicator than number of students educated.) While improvements in these areas should not be viewed simply as the outcome of the health (medical) or education system, because of impacts of other changes in the economy and society (health, as we have noted, is affected by diet and smoking), they can be viewed as measures of societal progress more generally.
112. Health outcomes are an arena in which distribution is as important as it is for income. There is, for instance, a wide dispersion of life expectancies. Several studies have shown that the relation between income distribution and life expectancy is sufficiently strong to produce significant associations in analyses of cross-sectional data and of data covering changes over time. The correlation almost disappears for the higher deciles (from the 7th to the 9th). The relationship is thus strongly non-linear, meaning that health is more responsive to changes in income among the least well-off⁴⁰.
113. For education outcomes, the best cross-country metrics of performance are provided by the Pisa test. But this test may not be the best for comparison across time within a country.
114. In the judiciary sector, an input-based measure of output may be especially misleading. The output of the sector should be some measure of (confidence in) justice. But a society characterized by a low level of trust between its citizens will have more courts,

38. David Cutler, *Your Money or Your Life, Strong Medicine for America's Health Care System*, Oxford University Press, 2005.

39. It is not easy to measure how our health care system is doing in extending life or reducing infant mortality. Cf. Angus Deaton, David Cutler and Adriana Lleras-Muney, "The Determinants of Mortality", *Journal of Economic Perspectives*, 2006.

40. R.G. Wilkinson, "Income Distribution and Life Expectancy", *British Medical Journal*, 18 January 1992.

more lawyers, and more conflicts. The consequent increase in GDP would not be a good measure of an increase in societal welfare.

115. While there are methodological disagreements about how to make the adjustments to quality or the measurement of government output, there is a broad consensus that some adjustments should be made, and even about the principles that should guide such adjustments. The disagreements arise in the practical implementation of these principles.

V. Economic Activity and Societal Well-Being

116. So far, we have focused on measures of *economic activity*, and even more narrowly, of market activity. But we are interested in measures of *well-being*. Here we will follow in the second tradition on national income measurement, launched notably by the work of Hicks and Samuelson, focusing on welfare measurement.
117. But going from a measure of market economic activity to a measure of economic well-being involves measurement, valuation, and conceptual problems (e.g. what aspects of economic activity should and should not be included).
118. While measures of economic activity focus on *production*, measures of well-being focus on household *income* and consumption. When we adjust for changes in prices, the market basket of consumption goods (reflected in the *consumer price index*) is different from the market basket of produced goods (reflected in what is called the GDP deflator). In addition, the market basket of consumption goods consumed by the rich and the poor, by the young and the old, by those in cities and those in rural areas, may differ markedly, and so we should have a different consumer price index for each of these groups. Reporting what happens to relative *nominal* incomes may not tell us all we want to know about relative *real* incomes. The most obvious example relates to the rapid rate of increase in urban housing and medical costs, with the result that the *real* income of both the elderly – at least those who are not sufficiently subsidized through an insurance scheme – and urban dwellers may be rising less rapidly than would be suggested by looking at their incomes and deflating by the “average” consumer price index.
119. When we look at an open economy, two important adjustments need to be made. We have already discussed one: income from domestic production may accrue to foreign companies and citizens, so measures of what is produced in the country (such as *Gross Domestic Product*) may differ from measures of the income of those within the country (*Gross National Product*, or GNP, which had been the dominant metric before 1990, the gross income (i.e. ignoring depreciation) of those living in the country. In the Report, we present refinements on this concept.)
120. The second adjustment relates to price changes. There may be marked differences between the prices of the goods produced by the country and the prices of the goods consumed by the country. This is most obvious in the case of oil exporters, who face

large fluctuations in the price of what they produce relative to the price of what they consume. The level of economic activity – the amount of oil produced – may change little, but the level of consumption that this oil may sustain may vary greatly.

121. For households, capital gains on the assets that they own can be an important part of their “income”. Yet national income statistics usually exclude these gains. The increases in stock and housing prices during the 1990s and the early years of this decade were an important contributor to individuals’ sense of well-being, and the capital losses in 2008 contributed significantly to individuals’ sense of a decrease in well-being. Year-to-year variations in wealth valuation should presumably be taken into account only if we are confident that markets value wealth correctly, an issue that we discuss more extensively in Section 1 above and Chapter 3 of the main Report, where we tackle the sustainability issue.
122. In addition to these two adjustments, several others are required in going from a measure of *economic activity* to a measure of individual *well-being*. Typically, we subtract out taxes to obtain a measure of (household) *disposable* income. Individuals do not directly get enjoyment out of the taxes they pay, but they do benefit from the services that government provides, and that is why it is important to add back in the (correct) value of public services. A few other adjustments are required to take account of: (a) intermediate goods; (b) security; (c) defensive expenditures; (d) leisure; (e) non-market activity; and (f) depreciation, resource depletion, and the degradation of the environment. Non market activity is discussed in section VI, and the issues posed by depreciation and depletion are postponed to the discussion of sustainability (Section VIII). The first three adjustments (combined with the other adjustments previously described) give us a measure that is sometimes referred to as adjusted disposable income⁴¹. The table below gives the real adjusted disposable income for France, the US and Finland. But in order to proceed to international comparisons we need to deflate nominal income or consumption. The problem is that both relative prices and consumption baskets differ across countries. Constructing appropriate price indices or *Purchasing Power Parities (PPPs)* have a long tradition in the OECD countries. Worldwide PPPs are periodically released by the World Bank. Their use avoids the many defects of comparisons based on nominal exchange rates. If, for example, the euro suddenly appreciates by 20% vis-à-vis the dollar, that does not mean that the adjusted disposable income of Europeans suddenly jumps by 20% relative to that of Americans. Of course, all the problems of construction of price indices apply to the construction of PPPs. Still, there is a broad consensus that PPPs provide the best bases of cross-country comparisons of standards of living.
123. The table below shows that the difference in real *adjusted* disposable income between France and the US—taking into account the benefits received from the government as well as the taxes taken away-- is much smaller (more than 20% smaller) than the difference in real disposable income between the two countries, just because the social transfers in kind received by the population are much greater in France. Indeed, the

41. Though the term “disposable” may not be fully appropriate, since individuals may not have “control” over the value of public services they receive. If the government has a balanced budget, and if the value of output is correctly valued by the value of inputs, then adding back the value of public services while subtracting out the value of taxes would leave the value of disposable income unchanged..

differences in real adjusted disposable income are less than the difference in real per capital GDP.

Table 3. Real income of households, comparisons across countries, 2005

		France	United States	Finland
Real disposable income* per capita	USD	19338	29448	14395
	USA=100	66	100	49
Real adjusted disposable income** per capita	USD	25378	32110	21771
	USA=100	79	100	68
Real GDP per capita (total economy)	USD	30519	41740	30460
	USA=100	73	100	73

*For private households and non-profit institutions serving households; net of depreciation; converted with PPPs for final consumption expenditure

**Equals disposable income corrected for social transfers in kind and converted with PPPs for actual individual consumption

Source: OECD Annual National Accounts.

a. Intermediate Goods

124. The general principle that intermediate goods should not be included as part of final output (income), because their value is already included in that of final goods, is well accepted and has already been discussed; yet it is not always clear what is an intermediate good. Most problematic are expenditures on the financial sector, where problems are compounded by the failure of other assumptions of the standard model. Consumers are not always well informed and markets are not always perfectly competitive, so that some expenditures – many would say a significant fraction of certain categories of expenditure – are really simply a transfer of income from households to the financial sector. The real issue here is the measurement of the “services” provided by that sector. These are measured by inputs augmented by profits. Again, such adjustments would not be so important, were it not for the importance that this sector has taken in some economies (accounting for a third or more of corporate profits).

b. Security

125. There is ample evidence that security – physical security, health security, and economic security – affects individuals’ sense of well-being. Changes in security represent some of the important changes in our societies. There has, in some cases, been improvement in the ability of markets to manage risks. Governments have taken on changing roles in managing risk. But neither in the case of the market or of government are the changes unambiguous in nature. Many countries have shifted from defined benefit to defined contribution pension schemes, and at the same time have cut back on publicly provided pensions. This has increased retirement insecurity. These are important changes in societal well-being, and yet the value of the “retirement security” provided traditionally by employers or by government is not captured in our system of national accounts.

126. Better (or worse) macro-economic management may result in less (or greater) economic volatility. Better (or worse) social insurance means that the risk borne by citizens as a result of this volatility may be less (or greater). It is possible but difficult to value the risk that individuals face – what they would pay to reduce or eliminate the risk – but our system of national accounts does not do this. For poor individuals, unable to access credit markets, with little savings beyond their retirement accounts, even a short period of unemployment may impose great hardship. But this problem is greater in the United States, where the loss of a job is typically associated with the loss of health insurance. In developing countries where the system of social protection is embryonic, economic stability is even more important, as fluctuations of output may have irreversible consequences on the destiny of important fractions of the population.
127. These are changes that have been a consequence of alterations in our *economic* system, which have large import for societal well-being, and yet our system of economic accounts does not reflect these changes. In assessing economic performance – both changes in time and differences across countries – we need to develop better metrics, at least for the extent of *economic* insecurity.
128. Economic security naturally falls within the remit of a system of economic accounts. But so too do other dimensions of security that affect well-being, such as our sense of physical security.

c. Defensive Expenditures

129. Increases in spending for security guards to protect property or prisons to incarcerate societal offenders are included in our system of national accounts, but the question is, should they be? These are not inconsequential – earlier, we noted the substantial spending in the United States on prisons. In the case of security guards employed by firms, these are intermediate expenditures – a cost of production. But more difficult problems (discussed in section 2.2) are raised by expenditures by households and government. This raises the fundamental questions posed earlier for other categories of public expenditure: We should be measuring the “output”, our sense of security; instead, we measure inputs. And, in assessing the economic system, we should be measuring the *net value added* by the economic (and perhaps social) system. If failings in the socio-economic system lead to more insecurity (crime), for instance, as a result of failures in education, then the increased expenditures on police and prisons is simply offsetting the deficiencies in our education system. Savings in one account are offset by increased expenditures in another. If, however, the increase in crime is a result of some exogenous force, unrelated to the socio-economic system, then the increased expenditures on prisons are increasing welfare *relative to what it would otherwise have been*. *Social progress* may be less than we would have hoped (because of the adverse changes in behaviour), and our metrics of social progress should reflect this fact; but there has been an increase in the “value added” of the economic system, in undoing these adverse changes, and our metrics of economic performance should reflect this as well.
130. As we noted earlier, there are many aspects of changes in expenditures in the process of modernization that fall within this category – the increased expenditures on environmental services to offset the costs of urbanization (the benefits of which, for instance, are reflected in the system of economic accounts in transportation costs that

are lower than they otherwise would have been). With the rise in the world's urban population from 1.6 billion to 3.3 billion in the last 30 years, these "urban defensive" expenditures may not be inconsiderable.⁴²

d. Leisure

131. A final extension of our *standard* model is to include leisure. Not including leisure in a measure of societal well-being would bias our measurement system to encourage the provision of goods and services – a bias that we should be particularly sensitive about as the world faces environmental problems posed by high levels of production. Standard theories suggest that societies might rationally take some of the benefits resulting from improvements in technology in the form of increased leisure – and it may be a symptom of a dysfunctional society if that is not the case. Yet a society's decision to enjoy more leisure as it becomes better-off would be counted against it in standard measures of GDP and the variants discussed so far. Some account of changes in leisure over time and differences in leisure need to be made, though again, as in so many of the other areas of non-market activity, there are problems of measurement and valuation, which are discussed at length in Chapter 1 of the main Report.
132. A particular problem is posed by involuntary unemployment – where individuals cannot work as much as they would like – or they enjoy more leisure and/or home work than they would like. Even if we had data with which we could ascertain what fraction of the additional time were devoted to leisure, we would have a problem of valuing it – because it is, in part, involuntary, the value is less than the (marginal) wage. (The problem of involuntary unemployment includes not just the "officially" unemployed, but also those working part-time involuntarily, because they could not find full-time employment.) But still, individuals can choose to engage in non-market activity (discussed in the next section), so that the value of leisure should be at least as great as the (marginal) imputed value of labour devoted to non-market activity.

VI. Non-Market Economic activity

133. GDP – and our modifications – began as a measurement of market activity (in the broadest sense of the term, including government). But much activity occurs outside the market, and this too has important implications for societal well-being.
134. Accordingly, the second set of changes focuses on going beyond market production to non-market production. As we noted, we already do this, in the case of owner-occupied housing. But home production is far more important. Here, even more than in the areas previously discussed, we encountered measurement and valuation problems – how

42. Nordhaus and Tobin's classic study that attempts to come up with a better measure of societal well-being included significant adjustments for these defensive expenditures. For example, they propose to subtract from total private consumption a number of components that do not contribute positively to welfare, in particular, commuting.

much of the activity has occurred, and how do we value these activities relative to, say, market activities. Some of the same problems referred to earlier arise in the measurement of improvements in productivity and changes in quality. As we broaden the scope of our definition, we have to rely increasingly on imputations, on educated guesses, and this reduces the confidence we have in the numbers, and thus their usefulness. This reliance on imputations is of particular concern because the magnitude of non-market activity (as conventionally measured) is so large. We are caught in a dilemma: because it is so large, to ignore it is to ignore important elements of human activity.

135. To be sure, we have much more data today to ascertain what individuals do when they are not engaged in work. Time-use studies allow us to break down more precisely the kinds of non-market activities individuals are engaged in. Problems of ascertaining which activities are “intermediate goods”, which are forms of “leisure”, and which are really non-market “production” abound. For instance, driving to work is part of the cost of market activity that is not reflected in market data. If a society spends more time traveling to work, it is not an increase in final output, but really reflects a reduction in productivity, almost surely reflected in a reduction in leisure (or perhaps some other market or non-market activity).
136. We may be more confident about measuring certain changes – shifts from home production to market production that occur broadly within society – and it would be wrong not to make note of these changes. Failing to do so may seriously bias our estimates of improvements in societal well-being. As we have already noted, as countries develop, there is often a shift from home to market production. The increases in market production may thus overstate increases in well-being. By the same token, policies that encourage market over non-market production distort the economy. Those engaged in non-market production often argue that not to include it in our measure of societal well-being is to devalue the important services provided by those who are engaged in these non-market activities. An important example may help make the point stronger. There is a serious omission in the valuation of home-produced goods – the value of breast milk. This is clearly within the System of National Accounts production boundary, is quantitatively non-trivial and also has important implications for public policy and child and maternal health⁴³.

VII. Societal Well-Being

137. The amendments we propose to traditional GDP accounting are incremental improvements. They allow us to come *closer* to what really matters for citizens, their well-being, or more generally, the quality of their lives. But while these are improvements, they leave out much that is important, partly because they focus on the

43. For a good discussion of this issue, cf. Julie Smith and Lindy Ingham, “Mother’s Milk and Measures of Economic Output”, *Feminist Economics*, Vol. 11, No. 1, March 2005, pp. 41 – 62.

present, and partly because well-being has a much broader content. It includes in effect the full range of factors that make life worth living, reaching beyond its material side. Fortunately, innovations in social science research provide us tools with which we can assess individuals' sense of their own well-being.

VII-1 Happiness, well-being and quality of life

138. Before beginning our discussion of broader metrics of societal well-being, some clarification is in order. In the public debate, the notions of Happiness, Well-being and Quality of life are used almost as if they were synonymous, but there are some distinctions that may prove useful. Without entering a semantic debate, it seemed to us from the outset that, while all of these notions have a subjective component, some might be better suited for measurement. In particular, there is a notion of “quality of life” (described more fully in the Report) that is more suited for measurement and for comparative assessment. In the discussion below, we will use this term interchangeably with the term “well-being”.
139. The central problem can be explained in the following way: all approaches to the measurement of the quality of life based on (marketed) resources (or on people's command over commodities) remain limited in important ways. First, many resources are not marketed. Second, many of the determinants of human well-being are not resources but aspects of people's life-circumstances. Lastly, resources are means that are transformed into well-being in ways that differ across people: those individuals with greater capacities for enjoyment or greater abilities for achievement in valuable domains of life may be better-off even if they command fewer economic resources.⁴⁴ These arguments, by themselves, suggest that metrics based solely on access to or command over resources are inadequate metrics for quality of life. Which other metric(s) should be used for assessing quality of life depends on the philosophical perspective taken.
140. Recent advances in research have led to new and credible measures. These measures, while not *replacing* conventional economic indicators, now have the potential to move from research to standard statistical practice. While some of these measures reflect structural conditions that are relatively invariant over time but that differ systematically across countries, others are more responsive to policies and more suitable for monitoring changes over short periods of time. Both types of indicator play important roles in measuring quality of life.

VII-2 Conceptual approaches to measuring quality of life

141. Three conceptual approaches have drawn the attention of the Commission as useful in thinking about the measurement of quality of life.
142. The first approach, developed in close connection with psychological research, is based on the notion of *subjective well-being*. A long philosophical tradition views individuals as the best judges of their own conditions. (This approach is closely linked to the

44. To link this discussion with that of the preceding section: some individuals may be more efficient in “home production,” others in market-production.

utilitarian tradition but has a broader appeal due to the strong presumption, in many streams of ancient and modern culture, that enabling people to be “happy” and “satisfied” with their life is a universal goal of human existence.)

143. The second approach is rooted in the notion of *capabilities*. This approach conceives a person’s life as a combination of various “doings and beings” (functionings) and of his or her freedom to choose among these functionings (capabilities). Some of these capabilities may be quite elementary, such as being adequately nourished and escaping premature mortality, while others may be more complex, such as having the literacy required to participate actively in political life. “In contrast with the utility-based lines of thinking, individual advantage is judged in the capability approach by a person’s capability to do things he or she has reason to value. A person’s advantage in terms of opportunities is judged to be lower than that of another if she has less capability – less real opportunity – to achieve those things that she has reason to value.⁴⁵” It is why the capability approach is focusing on freedom.
144. The third approach, developed within the tradition of welfare economics, weights the various non-monetary dimensions of quality of life (beyond the goods and services that are traded in markets) in a way that respects people’s preferences. Welfare economics has traditionally relied on the notion of “willingness-to-pay” to extend the scope of monetary measures to non-market aspects of life. But evaluations based on total willingness-to-pay may disproportionately reflect the preferences of those who are better-off in society. The “fair allocations approach”⁴⁶ tries to overcome this weakness of traditional welfare economics by explicitly referring to equity criteria. We use that terminology below.
145. These approaches have obvious differences, but also similarities. All these approaches point to the importance of a number of features for quality of life that go beyond command over resources. Measuring these aspects of well-being or quality of life requires the use of types of data (i.e. responses to questionnaires and non-market observations of personal states) that are not captured by market transactions.

VII-3 Subjective measures of quality of life

146. In recent years, much research has focused on what people value and how they act in real life. A significant part of this research has been undertaken by psychologists and economists based on subjective data on people’s reported or experienced well-being.
147. Subjective measures have always been part of the traditional tool-kit of economists and statisticians, as many features of our economies and societies (e.g. unemployment) are measured through people’s responses to a standard set of questions. (Not working is an objective statistic; but to be unemployed, one has to be “not working, but seeking employment”; seeking employment is typically a subjective statement, not based on particular actions, such as filling out an application form.) The specific feature of the

45. Amartya Sen, *The Idea of Justice*, *op.cit.*

46. Cf. Kolm S.C. 1972, *Justice et équité*, Paris: Ed. du CNRS. Translated as *Justice and Equity*, Cambridge, Mass.: MIT Press, 2000. Cf. also the survey on the theory of fair allocations by Thomson W., H. Varian 1985, “Theories of justice based on symmetry”, in Hurwicz, Schmeidler, Sonnenschein (eds.): *Social Goals and Social Organizations* (pp. 261–289). Cambridge, UK: Cambridge University Press.

subjective measures of quality of life discussed here is that what people report about their own conditions has no obvious objective counterpart: we can compare “perceived” and “actual” inflation, for example, but only respondents can provide information on their own subjective states and values. Despite this feature, a rich literature on these subjective measures concludes that the results of these studies are replicable and they help predict people’s behaviour.

148. Subjective approaches distinguish between the *dimensions* of quality of life (i.e. people’s subjective experiences and their evaluations) and the objective *factors* shaping them. They focus on the subjective dimensions of quality of life, which encompass several aspects. The most important of these are: a) people’s evaluations, either of their lives as a whole or of life’s various domains, such as family, work and financial conditions; b) people’s actual feelings, such as pain, worry and anger, or pleasure, pride and respect. Within this second broad category of people’s feelings, the research on subjective well-being distinguishes between positive and negative affects, as both may characterize the experience of each person.
149. All these aspects of subjective well-being (cognitive evaluations, positive and negative affects) should be measured separately to get a satisfactory appreciation of people’s lives. Subjective measures of people’s life-evaluations and affects provide measures of quality of life that can be monitored over time; some of these measures can also be compared across countries in reliable ways. The interest of these measures is that they provide information beyond that conveyed by income. For example, younger and older people in most developed countries report higher evaluations of their lives than prime-age people, a pattern that contrasts sharply with that for income.
150. One point where various subjective measures of people’s well-being agree is that unemployment has a very adverse effect on people’s quality of life. People who become unemployed report lower life-evaluations, even after controlling for their lower income. The adverse effects persist over time. The unemployed also report higher prevalence of various negative affects (sadness, stress and pain) and lower levels of positive ones (joy). One may also suspect that the adverse effects of unemployment are felt even by those who are not themselves unemployed, especially in societies where there is high unemployment. These subjective measures suggest that the costs of unemployment exceed the income-loss suffered by those who lose their jobs, reflecting the existence of non-pecuniary effects among the unemployed, and of fears and anxieties generated by unemployment in the rest of society.
151. These important advances in the measurement of subjective well-being are largely the result of the efforts of individual researchers and commercial data providers. But these data remain limited in terms of the statistical inferences that they allow. National statistical systems need to build on these efforts, and incorporate questions on various aspects of subjective well-being in their standard surveys.

VII-4 Objective features shaping quality of life

152. Both the capability and the fair allocation approaches give prominence to the objective conditions and opportunities available to people. Both of these conceptual approaches regard an expansion of people’s opportunities in these domains as intrinsically

important for people's lives. Future research should help clarify the role played by these objective features in subjective well-being, and might contribute to the critical task of an assessment of the relative weights to be associated with different factors.

153. The range of objective features to be considered in any assessment of quality of life will depend on the purpose of the exercise. In general, all these measures highlight that how societies are organized makes a difference for people's quality of life, and that their influences are not all captured by conventional measures of the associated economic resources.

a. Health

154. *Health* is a basic feature shaping both the length and the quality of people's lives. Its assessment requires good measures of both mortality and morbidity. Data gaps remain significant in both fields. Mortality statistics by age and gender document the risk of death confronting people and are used to calculate the expected length of a person's life. These indicators are today available in all developed countries but remain limited in large parts of the developing world, in particular for adults, and this limits the possibility of monitoring progress in achieving the UN Millennium Development Goals.
155. The state of advancement is far more limited for statistics on morbidity, a situation that has led to long-standing disagreements on whether or not declines in mortality have been matched by similar declines in morbidity. Existing measures of morbidity rest on a variety of sources: records of peoples' height and weight; diagnostics from health professionals; registers for specific diseases; and self-reports drawn for censuses and surveys. The variety of measures and underlying data is inevitable given the many manifestations of poor health, but is also a real obstacle to comparing countries and monitoring changes in people's health status over time. Measures are even sparser when moving from physical to mental disorders, despite evidence that these affect (at least in mild forms) a large share of people, that most of these disorders go untreated, and that their incidence has been increasing in some countries.
156. The variety of dimensions of people's health has led to several attempts to define a summary measure that combines both mortality and morbidity. Several combined indices of people's health exist, but none currently commands universal agreement.
157. There are large differences in health outcomes as well as in access to those factors (including resources) that affect health outcomes. National statistics that simply describe average outcomes accordingly provide an inadequate description of a country's health status, and are an incomplete basis for making comparisons over time or across countries, or for assessing either the extent of social progress or economic performance even in this limited dimension.

b. Education

158. A long tradition of economic research has stressed the importance of education in providing the skills and competencies that underpin economic production. But education (like health) matters for quality of life independently of its effects on people's earnings and productivity. Education is strongly associated with people's life-

evaluations even after controlling for their higher income. Further, better-educated people typically have better health status, lower unemployment, more social connections, and greater engagement in civic and political life. The consensus is that education brings a range of returns (monetary and non-monetary) that benefit both the person investing in education and the community in which they live. Measuring the size of these wider benefits of education is an important research priority.

159. Available educational indicators cover a broad range of fields. Some refer to inputs (e.g. school enrolment, educational expenditures and school resources), others refer to throughputs and outputs (e.g. graduation rates, number of completed years of schooling, standardized test measures of people's achievements in terms of literacy and numeracy or other cognitive skills). Which of these indicators is more relevant depends on the stage of development of each country and on the goal of the evaluation exercise. The available indicators highlight large differences across countries, with various different educational indicators sometimes highlighting contrasting patterns. Within countries, measures of inequalities in learning outcomes are especially important for youth at the bottom of the achievement scale and who are at risk of poverty and exclusion from well-paid and rewarding jobs in adult life. As education is an important predictor of many dimensions of people's lives, all social surveys should systematically include information on the learning experiences of respondents and of their parents, as well as information on other features that would enable a better assessment of the role of education in shaping the quality of life.

c. The balance of time

160. How people spend their time, and the features of people's personal activities, matter for quality of life irrespective of the income that they generate. They affect people's subjective well-being. The main activities discussed here are paid work, commuting, unpaid work and leisure time. Housing, although not representing an activity *per se*, is also discussed because it is the setting for a number of personal activities and because of its importance for the quality of life.
161. Paid work matters for quality of life not just because it generates income with which other material goods are purchased, but also because it provides identity to people and opportunities to socialize with others. Not all jobs are equally valuable in this respect, and this underscores the importance of collecting more systematic information on the quality of paid work, as is being done in the context of ongoing research on "decent work" pursued by a number of international organizations. Some national surveys provide information on many aspects of the distance between work and decent work, such as precarious employment, gender gaps in employment and wages, discrimination in the workplace, etc. Their practical use is, however, limited by their small sample size and by differences in the survey design across countries.
162. Commuting time is also a key feature of the quality of work, and its monitoring requires information on the number of hours spent travelling to and from work, as well as on the accessibility, quality and affordability of public transportation systems.
163. Unpaid domestic work, such as shopping and the care of children and other household members, is important from the perspective of assessing both the total amount of household services produced and how family chores are distributed between men and

women. As we noted earlier, changes in technology may affect not only the productivity of domestic work, but also the quality of the work experience. Certain activities, such as cooking, depending on how they are conducted and the circumstances under which they are conducted, may contribute either negatively or positively to a sense of well-being.

164. A long tradition of research has emphasized the importance of leisure-time for quality of life. This points to the importance of developing indicators of both its quantity (number of hours) and quality.
165. Finally, despite the importance of housing for a variety of social outcomes (such as children's education), no core set of housing indicators currently exists for international comparisons: remedying this situation would require better information on the number of people who are homeless or living in emergency shelters, and on housing quality (e.g. in terms of overcrowding and the availability of local services).
166. In several cases, suitable indicators in these various fields already exist, and the challenge is to improve upon what has been achieved in the past and to systemize the collection and analysis of data so that meaningful comparisons can be made across countries and over time. In other areas, however, existing measures remain seriously deficient, and progress requires investment in new statistical capacity. A case in point, cutting across all the activities described above, is that of measuring how people spend their time and their subjective experiences with various activities. Time is the natural metric for comparing personal activities, and one priority should be to develop measurement tools grounded on clear definitions and based on surveys with a consistent design, representative of patterns over a full year, and undertaken with sufficient regularity (all being requirements that are not often met). These surveys should provide information about both the amount of time spent in various activities and the enjoyment that they provide.

VII-5 Political voice and governance

167. Political voice is an integral dimension of the quality of life. Intrinsically, the ability to participate as full citizens, to have a say in the framing of policies, to dissent without fear and to speak up against wrong are essential freedoms and capabilities.
168. The opportunities for expression and the degree of responsiveness of the political system depend on the institutional features of each country, such as the presence of a functioning democracy, universal suffrage, a free media, and civil society organizations. This also depends on some key aspects of governance, such as legislative guarantees and the rule of law. Legal guarantees require effective implementation, which depends on how various institutions (e.g. the police, the judiciary and various administrative services) function, whether they are free from corruption, political interference and social prejudice, and whether they can be held accountable for their decisions.
169. Comparisons based on existing indicators of political voice, legislative guarantees and the rule of law highlight vast differences between countries, especially between those with a long history of democratic functioning and those that have moved from authoritarian to democratic regimes only more recently and that have not yet established the full range of freedom and rights. Even in the developed world, however, low trust in

public institutions and declining political participation point to a growing gap in how citizens and political elites perceive the functioning of democratic institutions.

170. Indicators of political voice and democratic governance should help evaluate the functioning of multiparty democracy and universal suffrage, the degree of decentralization in government decisions and the sense of participation have in the decisions being made, the presence of a free media and various freedoms (e.g. to form and join civil organizations, trade unions and professional bodies, or to participate in civic and social activities). Relevant indicators should cover the rights embedded in constitutions, laws, international covenants on human rights and basic freedoms, as well as the functioning of the judicial system. Indicators of many of these aspects are typically compiled by bodies outside the boundaries of national statistical systems and are mainly based on the opinion of experts. These indicators need to be complemented, and in some cases replaced, by surveys on citizens' own perceptions about the functioning of political, legal and executive institutions, the difficulties they face in accessing them, and the trust that they place in them. Such surveys also need to capture inequalities in access to these institutions across socio-economic groups.

VII-6. Social connections

171. Social connections improve quality of life in a variety of ways. The effects are both direct and indirect. People with more social connections report higher life-evaluations, and for many the most pleasurable personal activities involve socializing. The benefits of social connections extend to people's health and to the probability of finding a job, as well as to various characteristics of the neighbourhood where people live (e.g. crime or the performance of local schools). These social connections are sometimes described as "social capital" to highlight the benefits (direct and indirect) that they bring.
172. The drivers of changes in people's social connections are not always well understood, and should be an important arena of research. Developments within markets as well as certain government policies may have reduced the ties of individuals with their community. What is clear is that a decline in these ties may negatively affect people's lives, even when certain of the "economic" functions associated with these social ties are taken up by market or government alternatives, and even when these alternatives increase the level of (measured) economic activity (as in the case where the informal surveillance of neighbours is replaced by salaried security guards, or when family "insurance" is replaced by market insurance).
173. Research on social connections has traditionally relied on proxy measures, such as the number of associations to which each person belongs, or the frequency of activities assumed to result from social connections (e.g. altruistic behaviour and voter turn-out). However, it is by now accepted that these are not good measures of social connections, and that reliable measures require surveys that inform about peoples' behaviours and activities. In recent years, a number of statistical offices (in the United Kingdom, Australia, Canada, Ireland, the Netherlands and, most recently, the United States) have started surveys that measure various forms of social connections, such as civic and political engagement, membership and voluntary work in various organizations, relationship with neighbours and family members etc. Similar surveys should be implemented elsewhere, based on questions and protocols that allow valid comparisons

across countries and over time. Progress should also be made in measuring additional dimensions of social connections (such as trust in others, social isolation, availability of informal support in case of need, engagement in the workplace and in religious activities, friendship across lines of race, religion and social class), building on the experience already accumulated by some countries in these fields.

VII-7. Environmental conditions

174. Environmental conditions are important not only for sustainability, but also because of their immediate impact on the quality of people's lives. First, they affect human health both directly (through air and water pollution, hazardous substances and noise) and indirectly (through climate change, transformations in the carbon and water cycles, biodiversity loss and natural disasters that affect the health of ecosystems). Secondly, people benefit from environmental services, such as access to clean water and recreation areas, and their rights in this field (including rights to access environmental information) have been increasingly recognized. Third, people value environmental amenities or disamenities, and these valuations affect their actual choices (e.g. of where to live).
175. Measuring the effects of environmental conditions on people's lives is, however, complex. These effects manifest themselves over different timescales, and their impacts vary depending on people's characteristics (e.g. where they live and work, their metabolic intake).
176. Much progress has been achieved in the last two decades in terms of measuring environmental conditions. However, from a quality-of-life perspective, existing indicators remain limited in important respects: for example, indicators of emissions mainly refer to the aggregate quantities of various pollutants, rather than to the share of people exposed to dangerous doses. Existing indicators should hence be supplemented by regular monitoring, for instance, the number of premature deaths from exposure to air pollution; the number of people lacking access to water services and to nature; or the number who are exposed to hazardous levels of noise and pollution, etc. In addition, more work is needed to relate changes in environmental indicators to changes in well-being.

VII-8. Personal insecurity

177. Assessing personal insecurity requires identifying those external factors that put at risk the physical integrity of each person: crimes, accidents, and natural disasters are some of the most obvious factors. While these elements account for only a minority of all deaths, and they are captured by general mortality statistics, the rationale for having specific measures of their frequency is that they can have a different emotional effect than deaths related to medical conditions. The importance of these sources of insecurity on people's subjective well-being is reflected both in their efforts to avoid physical insecurity and the large impacts of bereavement on subjective well-being.
178. Manifestations of personal insecurity affect quality of life for a large number of people, with even larger numbers reporting fears of being a victim of a physical aggression. The most remarkable feature of these reports on subjective fears is how little they are related

to experienced victimization: countries with a higher share of people reporting fear of crime do not have a higher frequency of experienced victimization while, within countries, older and richer people feel more unsafe than younger and poorer people, despite being less likely to be a victim of crime.

179. These patterns highlight the importance of developing more regular and reliable measures of personal security to orient public discussions. Victimization surveys are an essential tool. But other tools need to be mobilized to assess the importance of other threats to personal security, such as domestic violence and violence in countries ravaged by conflicts and wars.⁴⁷

VII-9. Economic insecurity

180. Uncertainty about the material conditions that may prevail in the future reflects the existence of a variety of risks, in particular associated with unemployment, illness, and old age. The realization of these risks has obvious negative consequences for the quality of life of the person affected.
181. Job loss can lead to economic insecurity when unemployment is recurrent or persistent, the replacement rate is low, and workers have to accept major cuts in pay, hours or both to find a new job. The consequences of job instability are both immediate (as replacement income is typically lower than the earnings on the previous job) and longer term (due to potential losses in wages when the person does find another job). While indicators of these consequences are available for some countries, cross-country comparisons are difficult, requiring special investments in this direction. Fears of job loss can have negative consequences for the quality of life of each worker (e.g. physical and mental illness, tensions in family life), for firms (e.g. adverse impacts on workers' motivation and productivity, lower identification with corporate objectives) and for society as a whole.
182. Illness can cause economic insecurity both directly, through the medical costs associated with it, and indirectly, through the loss of income due to inability to work. One indicator of economic insecurity due to illness is provided by the share of people without health insurance. However, health insurance can cover different contingencies, and even insured people may incur high out-of-pocket health expenses in the event of illness.
183. Old-age is not a risk *per se*, but it can still imply economic insecurity due to uncertainty about the needs and resources that will be available after having withdrawn from the labour market. Two types of risk, in particular, are important. The first is the risk of inadequate resources during retirement, due to insufficient pension payments in the future. Even when there is no risk associated with the pension amounts, there may be uncertainty about the adequacy of pension payments when they are not indexed to inflation. And even when they are inflation-indexed, if well-being is affected by *relative* income then when incomes of workers are rising rapidly, the aged may feel a risk of relative deprivation. The second is the risk of volatility in pension payments: while all

47. This is an example where the collection of objective data might actually affect individual perceptions, and therefore there sense of well-being.

retirement-income systems are exposed to *some* types of risk, the greater role of the private sector in financing old-age pensions and the consequent shifting of risks from government and firms towards individuals has led to increasing the volatility of pension payments.

184. A comprehensive measure of economic insecurity would ideally account for both the frequency of each risk and its consequences, and some attempts in this direction have been made. A further problem is that of aggregating across the various risks that shape economic insecurity, as the indicators that describe these risks lack a common metric to assess their severity.

VII-9. On the interdependence of the various dimensions of quality of life

185. Most of the measurement challenges described above are specific to each feature of quality of life, and the Commission has only hinted at some of the work required, leaving it to future work in each field to detail a concrete action plan. Other challenges are, however, cross-cutting and unlikely to be picked up through initiatives separately undertaken in each field. Three of these issues deserve special attention.

Assessing links across quality of life dimensions

186. The first cross-cutting challenge is to better assess the *relationship* between the various dimensions of quality of life. Some of the most important policy questions for quality of life relate to how developments in one area (e.g. education) affect developments in others (e.g. health status, political voice and social connections), and how developments in all fields are related to those in income. In effect, ignoring the cumulative effects of multiple disadvantages will provide an inaccurate assessment of the extent of certain social problems, possibly leading to sub-optimal policies: for example, if the loss in quality of life of being both poor and sick far exceeds the sum of the two separate effects, governments may need to target their interventions more on those who cumulate these disadvantages.
187. Assessing these links across various dimensions of quality of life will not be easy, as statistical systems continue to be highly segmented across disciplines. But progress can be achieved by developing information on the “joint distribution” of the most salient features of quality of life (such as health status, education, political voice, etc.). across all residents in a country. Concrete steps in this direction could be accomplished by including in all surveys a few standard questions that allow classifying respondents based on a limited set of characteristics that describe their conditions in a broad range of fields.

b. Inequalities in quality of life

188. The second cross-cutting challenge is to develop indicators of quality of life that provide information about the *inequalities* in individual conditions in the various dimensions of quality of life rather than just about the *average* conditions in each country. When combined with work on the linkages across “conditions” we can obtain better assessments of the extent of inequalities in quality of life.

189. While established methodologies and data sources could allow measuring inequalities in the distribution of economic resources in a fairly reliable way, the situation is much less satisfactory with respect to non-monetary dimensions of quality of life.
190. The problems, however, go deeper than developing suitable measures. There are many types of inequalities, and each of them is significant in itself: this suggests that we should avoid the presumption that one of them will always encompass all others. At the same time, because of the links among dimensions, various types of inequalities may strengthen each other. Gender disparities, for example, while pervasive in most countries and groups, are typically much larger for households with lower socio-economic status: the combined effect of gender and socio-economic inequality is often to exclude young women from poor households from attending school and getting rewarding jobs, denying them possibilities of self-expression and political voice, and exposing them to hazards that put at risk their health.
191. There exists a wealth of interesting research on inequality and deprivation in dimensions other than material resources, and much has been written on the role of composite indicators. For instance, Mackenbach (2006) points out that between-group difference are especially important in the case of health. In all countries around the world, people with lower education, lower income and from lower occupational classes tend to die at younger ages and to have, within their shorter lifetimes, a higher prevalence of different health problems. In European countries, the mortality rates of less-educated men are, on average, 50% higher than those of more-educated men, while the difference is 30% for women⁴⁸. A 2006 OECD study on education highlights the magnitude of the differences in within education: the test scores (in science) of students aged 15 in France were around 146 points higher at the top quarter of the achievement scale than at the bottom quarter, a difference equivalent to almost 4 years of schooling⁴⁹. With respect to composite indicators, detailed surveys of the many initiatives in this area are provided by Afsa et al. (2008) and Gadrey and Jany-Catrice (2007)⁵⁰.
192. Building on this work, it is critical that inequalities be assessed in a comprehensive way, by looking at differences in quality of life across people, groups and generations. Further, as people can be classified according to different criteria, each with some relevance for the quality of people's lives, inequalities should be measured and documented for a plurality of groups. Appropriate surveys should be developed to assess the complementarities between the various types of inequalities, and to identify their underlying causes.

c. Aggregating across quality-of-life dimensions

193. The third challenge to quality-of-life research is that of *aggregating* the rich array of measures in a parsimonious way. The search for a scalar measure of quality of life is

48. Mackenbach, J. P., "Health Inequalities: Europe in Profile", *UK Presidency of the EU*, February 2006.

49. OECD, "Programme for International Student Assessment" (2006).

50. Afsa, C., Blanchet, D., Marcus, V., Mira d'Ercole, M., Pionnier, P.A., Ranuzzi, G., Rioux, L. and Schreyer, P. (2008), "Survey of existing approaches to measuring socio-economic progress", *background paper for the first meeting of the CMEPSP*, Paris, April 2008. Gadrey, J. and Jany-Catrice, F. (2007), "Les nouveaux indicateurs de richesse", 2nd edition, *Repères-La Découverte*, Paris.

often perceived as the single most important challenge in this field. While this emphasis is partly misplaced – there are still important challenges in the construction of each of the indices, and the information content of any aggregate index will always reflect the quality of the measures used in its construction – there are important reasons to attempt to construct aggregate metrics.

194. Traditionally, the most common response to these demands for *parsimony* has been to aggregate a number of indicators (suitably selected and scaled) of *average performance* in various fields at the country-level. The best example of this approach is the *Human Development Index*. This measure has played an important information and communication role, leading to country-rankings that differ significantly from those based on per-capita GDP, especially for some less developed countries. However, all choices on the weights used to construct this (and other similar indices) are controversial – they are either arbitrary or reflect value judgments on which there is not broad consensus. More fundamentally, many such indices (including the HDI) ignore the distribution of individual conditions within each country and the linkages among conditions. As a result, the combined country-wide index will not change if average performance in each domain remains the same while the correlation of individual conditions across domains declines.
195. Several aggregate measures of quality of life are possible, reflecting the various conceptual approaches discussed. Some of these measures are already being used (e.g. average levels of life satisfaction⁵¹ for a country as a whole). Others could be implemented if national statistical systems made the necessary investment to collect and make public the type of data needed to allow their computation.

Concluding remarks

196. It is clear that individual's sense of well-being is affected not just by their material goods (appropriately measured and valued), but by many other factors, some of which we have called attention to earlier in this overview. For some of these, there may exist objective metrics, but for others, subjective assessments may provide the best approach to measurement. Individuals may, for instance, be affected by their sense of security, and by their bonds with others. But even the seemingly non-economic factors are affected by economic structures. Reforms in the workplace may lead to increased *market* efficiency, but lower worker job satisfaction, and therefore a reduction in their sense of well-being. One of the criticisms of globalization (in the way it has proceeded) is that it has contributed to the weakening of a sense of community.
197. Moreover, it is important to understand more fully the links between various measures of market and non-market activity and of leisure and these *quality-of-life metrics*. It is not clear how well the relative marginal contributions of various goods and services to well-being correspond to market prices.
198. This work is just at its beginning stage, and yet the results obtained so far are extremely promising. Replicable measurement is possible.

51. Data on life-evaluations have been collected in several representative surveys. In several cases (e.g. various waves of the *World Values Survey*), these measures are based on qualitative responses, such as feeling “quite” or “fairly” happy with one's life, or on other scales that assess life satisfaction.

VIII. Sustainability

199. Until now we have been concerned with the measurement of the “present”, whether in quantitative or in qualitative terms – how, for instance, certain market or non-market activities today affect the sense of well-being today. But our Commission aims at going beyond the measurement of individual and societal well-being today, moving the horizon towards the long run. What is missing in our information set to be tranquil about the future? The answer is obviously our capacity to assess the sustainability of our present situation and more generally of the economic, social, political and environmental context in which we live.

VIII-1 What does sustainability mean?

200. Sustainability is a concept which by definition has a long-term dimension. Sustainability refers, in a broad sense, to the notion of the durability and stability of dynamic processes in the long run. Concern for the sustainability of economic development can be traced back at least as far as Malthus’ famous population principle. Long before the Club of Rome, the classics were predicting that the world will be characterized in the long run by stagnation (zero growth), and though some of the reasons for their gloom differed, these analyses were generally based on the fact of the finiteness of the world.
201. More recently, the Brundtland Commission (1987) has popularized the definition of sustainable development as “development that meets the needs of the present generations without compromising the ability of the future generations to meet their own needs”. This definition implicitly mixes social, environmental and economic components of present and future well-being. But while the Brundtland report had the great merit of underlining from the outset that sustainability was a multidimensional phenomenon, it did not offer any precise evaluation of these components. It was, in this sense, both too general and too abstract.
202. The plea for sustainable development has been sounded so repeatedly in political discourse that sometimes it seems to have lost credibility, not only because actions did not follow promises, but because the meaning of the concept was so fuzzy that everyone was free to understand it according to their own interpretation.
203. At the micro level, sustainability means that individuals and/or families think that the future for them and their children and grandchildren will be better than the past, or at least not worse than the present. This positive expectation not only contributes to the current generation’s sense of well-being, but also has consequences for the nature of society. It has been argued, for instance, that when most families hold such hopes, society becomes more tolerant, fairer, and more supportive of democracy⁵². The opposite occurs when an economic evolution proves to be unsustainable. Fear of the future generally leads to the development of protectionist attitudes, “beggar my neighbour” policies, and anti-immigrant sentiments. None of this should come as a surprise: precariousness, which is the converse of sustainability, is an impediment to a

52. Cf. Benjamin Friedman, *The Moral Consequences of Economic Growth*, Alfred A. Knopf, 2005.

good life. Hence social progress, which is the thing people really care about, goes hand in hand with sustainability. This is why it is essential to attempt to develop a metric to account for it, however imperfect such a metric may be.

204. Whatever measure we devise, we have to recognize that it will be grounded on our present imperfect knowledge of the *future*. Even high levels of consumption of natural resources might be sustainable, if there is rapid enough technological progress.
205. Even if there are doubts about the precise or best measure, doubt does not justify paralysis, and for the reasons set out above, there is some urgency to come up with a measure or a set of measures of sustainability.
206. The literature gives many different definitions of sustainability. A very simple but natural formulation is the following: Assume that C is the current state of consumption, and $W(C)$ is the wealth that is required to sustain that level of consumption forever. Then consumption is weakly sustainable if the wealth bequeathed to the next generation (time period) exceeds $W(C)$. There is a more demanding concept of sustainability that focuses not just on the level of consumption but also on its rate of growth. If C is a consumption path – defined not just by the level of consumption but by its rate of growth and if current resources are not sufficient to sustain that consumption path, then the consumption path is not sustainable. The worry today is that, because of the finiteness of our world, even sustainability in the weak sense might not be assured. This worry is grounded on the concern that while the stock of some resources are increasing (physical capital), these cannot compensate for the diminution of the stock of others—in particular key environmental resources.
207. Much of the literature focuses the degree of substitutability between the different components of wealth. The “weak” approach to sustainability considers that increases in physical capital and/or human capital can compensate for losses in natural wealth, i.e. that there always exists a positive degree of substitutability. This allows a global assessment of sustainability using a single measure of “wealth”. “Weak” here means that the requirement for sustainability is not too demanding, as it implies that a sufficiently high pace of technical progress or savings is all that is needed to ensure sustainability. The “strong” approach argues that sustainability requires separately maintaining the quantity or quality of many different environmental resources. It is motivated by a concern that we may not be able, for instance, to deal with the worsening of the atmosphere by increases in the amounts of physical capital, or more generally, that even if we could, there is a high degree of uncertainty about the trade-offs, so much so that it is worth focusing on each of the key resources separately. Following this up therefore requires large sets of separate statistics, each pertaining to one particular subdomain of global sustainability.

VIII-2 On some complexities in measuring sustainability

208. 201. *In principle*, the most direct (and comprehensive) way to measure well-being in the long run is “simply” to add future consumption to current consumption – weighed by the appropriate marginal valuations (just as we add up the consumption of oranges and apples), to get an appropriate long-run metric. If we followed that approach, we would focus on “consumption” in each period, not on broader measures of national output,

which include investment. Investment is an intermediate good – an input into the production of consumption in future years. In this approach, the critical decision is the weight to attach to future consumption, that is the “discount rate”, which measures the extent to which future consumption gets less weight than current consumption.⁵³, a kind of measure of the depreciation of the future.

209. 202. In this approach, if future feasible consumption is less than the current level, it is obvious that the current level is not sustainable. If it exceeds the current level, it is obvious that the level is sustainable.
210. There are some problems with this approach, quite apart from the fact that, as we have seen above, the standard statistical apparatus may give a partial measure of total consumption. A critical difference between consumption today and that in future years is that we know today’s consumption, while we can only form expectations concerning future consumption. Economists traditionally have a way around this problem: they turn to the market to assign values, just as they do contemporaneously; the value of an asset is the expected present discounted value of the dividends (returns) that it yields. Thus, by looking at asset values, we can estimate the value of (expected) future consumption – or at least that part of future consumption that will be generated by our land, physical and human capital and other resources.
211. Recent events have cast doubt (for those who previously did not have such doubts) on the extent to which the current market prices of assets provide good assessments.
212. Risk itself poses considerable complexity in interpreting what we should mean by sustainability. Assume that there is uncertainty about the future rate of technological progress. Should sustainability require that we have a 99% probability of continuing current levels of consumption? A 99.9% probability? If we impose a very high probability, we would force the current generation to sacrifice current consumption for future generations, who almost surely will be much better-off.⁵⁴ Again, we often rely on market risk assessments. But the current crisis has again highlighted the risk of doing so. The market may overestimate the ease with which physical capital can substitute for the loss of natural capital, and as a result, may assign too low of a probability to the risk that current levels of consumption are not sustainable. This is belabouring the obvious, as markets are certainly not as far-sighted as to account for what may happen in two or three generations! Can we do better than the market? That is almost a philosophical question which can be answered, considering the deep uncertainty surrounding the issue, only by political decisions based on some form of consensus. Besides, what is of concern is the well-being of future generations. This is an issue of intertemporal equity, and interest rates are of limited relevance in providing guidance on this issue.⁵⁵ Making matters worse is that markets for key environmental variables (like “carbon”) do not exist, and the dramatic change in prices that might occur were appropriate social prices

53. The appropriate discount rate has been one of the critical issues under debate in the environmental literature. It is more relevant for discussions of policy than the issues upon which the Commission focuses. Those who believe that we should discount future consumption at a high rate typically are simply not concerned with sustainability: the loss of well being of future generations is outweighed by the gains to the present generation.

54. Further complexities are added to interdependencies: in the circumstances in which climate change turns out to be worse than expected, much of the capital stock may turn out to be in the wrong place (e.g. under water).

55. There are some extreme cases—infinately lived individuals, or finitely lived individuals with perfect altruism—in which market prices would be of relevance; but these assumptions have little to do with the world in which we live.

imposed might have large impacts on current values of interest rates and the relevant risk premia.

213. Still, it is important for any society to form an assessment, no matter how imperfect, about whether its current consumption or well-being is sustainable, and whether this is coming at the expense of future generations.⁵⁶ We can ascertain whether a society's wealth is increasing or decreasing (per capita). If (appropriately measured) it is increasing, then presumably society can do in the future whatever it did today, i.e. it can sustain its per-capita income. But we need a *comprehensive measure* of wealth, and we need to be sure that the valuations are correct. A comprehensive measure obviously includes measures of physical capital, human capital and natural capital (including the environment). Changes in capital include those arising from investment in plant and equipment, education, the depletion of natural resources, depreciation of physical capital, and environmental degradation.
214. Accordingly, we believe that a good system of national accounts should report both a level of consumption and a comprehensive measure of the change in wealth $\{C, \Delta W\}$. The task is difficult, because many components of wealth are not measured at all (i.e. human capital) or are often ill-measured.
215. The current crisis highlights the problem of relying on market prices for the valuation of wealth. Net wealth as measured was increasing in periods prior to the crisis, but that was the consequence of a market failure. Private debt was increasing, but according to the market, asset prices were increasing at a faster pace. A reliance on market prices would have led to the conclusion that the high levels of pre-crisis consumption in many countries were sustainable. The revaluation of wealth that occurred subsequently showed that they were not. The ongoing controversy over mark to market accounting shows that today, even ardent advocates of markets have lost faith in their valuations, at least in the short run. The difficulty is finding an alternative.
216. Market prices fail to provide accurate and reliable estimates of the expected present discounted value of future consumption (returns) not just because of the irrational exuberance and pessimism to which markets are prone. Market prices may be highly distorted, for instance, at one time putting too low a price on risk, and at another too high a price.
217. The most serious distortions in market prices arise out of the failure to price scarce environmental resources. The market today assigns no or a low price to carbon emissions, and yet there is a broad scientific consensus that there is a real scarcity value. The world is rapidly using up the available global "carbon space", but those who are using it are not being charged. The result is that the prices of all goods and services that make use of this carbon space – essentially *all* goods and services – are being distorted.

56. What is relevant, of course, is not just the level of aggregate consumption, but also particular policies. There is, for instance, concern that America's health care system is *unsustainable*, in the sense that it seems to require an ever-increasing share of GDP. Some economists use forecasts of future prices to estimate the necessary taxes that would have to be levied to sustain it, or, in the absence of additional taxes, the magnitude of the debt. One has to be careful, however, using such naïve extrapolations. *Something* will have to adjust – the policies are correctly identified as "non-sustainable". But it is probably wrong to infer that expenditures will simply increase in the way forecast. There are limits to the share of GDP that a society will be willing to devote to medical care.

218. We need to take a comprehensive approach to the measurement of wealth, including all assets and all liabilities. Liabilities include the costs of future waste disposal and repairing environmental damage.
219. These adjustments for changes in wealth are even more important when it comes to resource-dependent countries. Their stock of resources is being depleted, and their wealth is increasing only if the proceeds are being reinvested, whether in physical, human or financial capital.
220. While national income accounts have long taken account of the depreciation of physical capital, there are serious problems in the measurement of depreciation (none of the standard methodologies correspond to the theoretical construct of true economic depreciation), and these problems may have become more serious, as we noted earlier, with increases in the pace of innovation and the consequence increases in the rate of obsolescence.
221. One commonly used measure of *net* national income entails simply adding up (with appropriate prices) consumption and the change in wealth. “Green” measures of GDP/NDP/NNP subtract from the standard measures the value of the depletion of natural resources and environmental degradation. While such extensions are a move in the right direction, they typically do not embrace a sufficiently comprehensive measure of wealth. Valuing environmental inputs into the economic system is the (relatively) easier step. Since these inputs are incorporated into products that are sold in the market place, it is possible (in principle) to use direct means to assign a value for them based on market principles. In contrast, as pollution emissions are outputs, there is no direct way to assign a value to them. All the indirect methods of valuation will depend to some extent on “what if” scenarios. Thus, translating valuations of degradation into adjustments to macro-economic aggregates takes us beyond the realm of *ex-post accounting* into a much more hypothetical situation. The very speculative nature of this sort of accounting explains the great discomfort and strong resistance among many accountants to this practice.
222. For a country that was on an unsustainable consumption path, an analysis employing a comprehensive and accurate measure of wealth would show a time series of decreasing net national income.⁵⁷
223. Three remarks are in order. Even if we agree that the level of present patterns of consumption is unsustainable, in view of the adverse effect on the environment, it does not necessarily imply that future generations should consume less. Such a statement would be true only if we assume that they will consume the same basket of goods. But they may well change their consumption habits – because their tastes evolve, because some new products appear and/or because of a change in relative prices due to the increasing scarcity of natural resources. Indeed, these changes in prices are likely to induce both innovations and changes in consumption patterns. Second, the price of a resource near to exhaustion rises to infinity only if this resource has no substitute. And

57. For the world as a whole, as the resource approaches zero, unless there are substitutes, the price would approach infinity; presumably, the ever-rising price would induce innovation and extreme conservation in the utilization of the resource, so that the pace of utilization of the resource would be greatly reduced. (Many individual countries have, of course, seen the exhaustion of particular resources, e.g. the UK’s supply of North Sea oil.)

that depends on a number of phenomena, including technical progress, which are impossible to ascertain today. Third, this analysis highlights the importance of risk – and the valuation of risk.

VIII-3. What do we want to measure?

224. There have been many attempts at building indicators of sustainable development. Some of them directly stem from an accounting framework (the Nordhaus-Tobin approach) and are therefore familiar to economists or accountants. Many of these have been based on the Samuelson-Hicks welfare approach. Some others are more specifically environmental and have become very popular amongst NGOs and environmentalists. Several statistical bodies or scholars have, in parallel, developed eclectic approaches that combine several dimensions of sustainability, under the form either of extensive dashboards or of so-called “composite” indicators that add up, in one way or another, various indicators to form an “index”.
225. The situation therefore appears to be characterized by an excess bounty, rather than by scarcity. But this is not necessarily good news, especially when different indicators provide diverging messages about the sustainability of different economic models or about the contributions of different countries to worldwide sustainability. This has been a source of perplexity for public opinion and for policy makers: which indicators should one focus upon, which should be priorities for further development? So far, a consensus has not developed for any of the existing approaches.
226. What do we really want to measure? Since the Brundtland Report, the notion of sustainable development has expanded to become an encompassing concept absorbing all dimensions of present and future economic, social and environmental well-being. The two first chapters of the Report have concentrated on the measure of current well-being. Here we are concerned with the sustainability of this well-being, or, what amounts to the same thing, the “sustainable” component of “sustainable development”. This question of durability can be expressed in the following terms: assuming we have been able to assess what is the current level of well-being, the question is to know whether the continuation of present trends allows (at least) the preservation of this current level of well-being. It is why we have asserted that social progress in the long run presupposes sustainability.
227. As a first guide for sorting out the many different approaches reviewed in our Report, we separate these two notions, assessing current well-being and assessing its durability or sustainability, because the two questions are analytically distinct.
228. Some dashboards of sustainable development have mixed together measurement of current well-being and of its sustainability. While dashboards are useful, we want to select a limited number of indicators – a “micro” dashboard – including some indicators specifically dedicated to the sustainability issue, based on a clear notion of what sustainability means.
229. Composite indicators based on such eclectic dashboards raise similar problems, with one additional one – the way in which their various items are weighted is arbitrary, with

consequences that are seldom made explicit.⁵⁸ Such aggregation procedures are sometimes presented as superior to the monetary aggregations that are used for building economic indices, precisely because they are not linked to any form of market valuation. Indeed, there are many reasons why market values cannot be trusted when addressing sustainability issues, and more specifically their environmental component. But monetary or not, an aggregation procedure always means putting relative values on the items that are introduced in the index. The problem is not that these weighting procedures are hidden, non-transparent or non-replicable: they are often very explicitly presented by the authors of the indices – and this is one of the merits of this literature. The problem is rather than their normative foundations or their societal implications are seldom made explicit or justified.

VIII-4 Measuring our legacy to the future

230. As we have seen, the capacity of future generations to have standards of well-being at least equal to ours depends upon whether we pass on to them sufficient amounts of all those assets that matter for well-being. Let's note W the "extended wealth" index that would be used to quantify this stock of resources. Measuring sustainability amounts to assessing whether this global stock (or particular components of it) is evolving positively or negatively, i.e. computing current rates of change dW (or dW_t). If this is negative, downward adjustments in consumption or well-being will be required sooner or later. This is exactly what one should understand by "non-sustainability".
231. In our view, such a formulation of the sustainability issue has a strong potential for providing the common language necessary for constructive debates between those coming from very different fields and perspectives. Just to take one example, it fully answers one of the longstanding objections made to GDP by environmentalists, i.e. the fact that ecological catastrophes can increase GDP through their implied impact on economic activity. In an extended wealth approach, an ecological catastrophe is registered as a destruction of capital. This accounts for the fact that it decreases sustainability by decreasing the resources available for generating well-being tomorrow. This can be avoided only if some action is taken to repair the corresponding damage, these actions being counted as positive investment.
232. Some environmentalists worry that the depletion of certain key resources that are *essential* cannot be compensated for by, say, more investment in physical capital. Economists respond that if that were the case, markets should reflect that scarcity by a very high price, which would induce conservation of that resource and the development of substitutes. Techno-optimists point out that since Malthus, there have been those forecasting doom – whether due to a scarcity of land or of some other resource. *So far*, this has not occurred. But environmentalists say that markets are short-sighted, and that we have repeatedly seen their inability to take into account future effects.
233. This criticism holds, however, only when the indicators are computed at fixed prices. But if a natural resource becomes crucial for human survival, and cannot be replaced by other produced assets, a correct implementation of the index would posit a relative price

58. In contrast, market based indices use relative prices, which, when markets are competitive, reflect individuals' marginal valuations.

for the natural resource that tends to infinity. In that case, no realistic change in other assets will be able to restore sustainability. Appropriately constructed indices will reflect the risk of such changes in relative prices.

234. Our Commission believes that, nonetheless, there are sufficient societal concerns about certain key aspects of the depletion and degradation of our natural resources that reporting what is happening to certain key physical stocks may be warranted.
235. A difficulty in discussions of sustainability using the extended wealth framework (as conventionally employed) is posed by the apparent contradiction between its message that we can easily compensate for the depletion of resources and the degradation of the environment and the many arguments that have been developed in favour of strong and rapid actions in several environmental domains, climatic change being the most emblematic of them. The power of compounding can give the feeling that there will always be sufficient money available to pay for unfavourable environmental impacts, and that current problems are therefore more important. Simplistic uses of monetary indices of sustainability (often incorrectly constructed, paying insufficient attention to future potential risks) may seemingly provide some support to such a bias in favour of the present. The approaches we recommend are designed to avoid such a bias; we believe that it is imperative to pay due attention to the problems arising from the depreciation of environmental assets that are essential to human well-being or even survival.

a. “True saving”

236. Measures of “true savings” represent more comprehensive attempts at measuring changes in wealth, using “correct” valuations. As we have emphasized, wealth includes everything that can be transferred through time, whether privately or collectively owned.
237. Interestingly, most approaches (including those described at length in Chapter 3 of the main Report) do not rely on market valuations, simply because the presumption that markets do not work for purposes of long-term intertemporal valuations is so strong. They do not work well, even in the case of standard marketed goods, but even less well in the area of the environment, which is the focus of so much of the concern about sustainability.
238. It is also why we argue for a pragmatic approach that combines limited-scope “synthetic” indicators (in that they aggregate over a wealth of items, like true savings) to measure overconsumption with a few, well-chosen physical measures. We have, however, reservations about very encompassing approaches – which, in our view, suffer from their excessive ambition, e.g. summarizing current well-being and its sustainability with a single number – as well as about those that focus on a long list of separate resources, implicitly assuming no substitutability among them or between them and other forms of capital.
239. We argue for an extended wealth framework. This requires information that is not often available. Moreover, there is often a high degree of uncertainty about the appropriate trade-offs (weights associated with the various components of an extended wealth measure). We thus argue in favour of a multi-dimensional approach to sustainability. We

emphasize the problems raised by technological and other uncertainties, as well as those raised by the international dimension of the sustainability issue. We suggest staying with a modest approach, i.e. focusing the monetary aggregation on items for which reasonable valuation techniques exist, such as physical capital, human capital and fossil resources. This should be complemented with a limited set of physical and other non-monetary indicators dealing with climate change and global warming, and biodiversity, as well as the social capital and “institutional assets” that we transmit to future generations.

b. Footprints

240. Although apparently quite different from “extended wealth” notions, various attempts at measuring sustainability through the use of “footprints” are, in fact, also inspired by the general approach of comparing the level of an existing stock with current consumption flows and their induced effects on some dimensions of the environment. In this sense, they may also be regarded as “wealth” measures in which the focus is exclusively on natural capital, and the valuation convention differs in that no market prices are explicitly used. The results are expressed not in terms of the conventional “money metrics” but in terms of some physical metric (such as the total use of carbon, or the land required to sustain consumption levels).
241. The Ecological Footprint (EF hereafter) measures “how much of the regenerative capacity of the biosphere is used by human activities (consumption)”. It does so by calculating the amount of biologically productive land and water area required to support a given population at its current level of consumption and resource utilization. A country’s Footprint (demand side) is the total area required to produce the food, fiber and timber that it consumes, absorb the waste that it generates, and provide areas for its infrastructures (built-up areas). On the supply side, “biocapacity is the productive capacity of the biosphere and its ability to provide a flux of biological resources and services useful to humanity”, as Moran, Wackernagel and their co-authors (2008) phrased it.
242. The results are well-known and rather striking: since the mid-1980s, humanity’s footprint has been larger than the planet’s carrying capacity, and in 2003 humanity’s total Footprint exceeded the Earth’s biocapacity by approximately 25 per cent (we would have needed an extra 25% Planet to meet our needs, to put it informally). While 1.8 global hectares per person are available world-wide, Europeans use 4.9 global hectares per person and North Americans use twice that amount, that is, much more than the actual bio-capacities of these two geographical zones.
243. This index may be regarded as an extended account approach, even if its results are not expressed in monetary terms. Indeed, this indicator shares with accounting approaches the idea of reducing heterogeneous elements to one common measurement unit (the global hectare, e.g. one hectare with productivity equal to the average productivity of the 11.2 billion bioproductive hectares on Earth). It assumes the substitutability at least of different forms of natural capital (demands on different natural capital goods are additively assessed in terms of land area), but implicitly rejects notions that physical or other forms of capital can substitute for natural capital. In fact, there is no virtue to any saving or accumulation: not only does any positive ecological surplus (biocapacity

exceeding ecological footprint) not entail an increase in some natural capital stock and hence an improvement of the productive capacity in the future, but investments in capital goods that might reduce future ecological footprints have an adverse effect because of the demands that they impose on today's ecological footprint.

244. The results of this and similar metrics are problematic for measuring a country's own sustainability – or would be relevant only in a world of autarky. Differences in resource endowments provide one of the main motivations of trade. The fact that densely populated (low biocapacity) countries like the Netherlands have ecological deficits whilst sparsely populated (high biocapacity) countries like Finland enjoy a surplus can be seen as part of a normal situation where the trade of goods is mutually beneficial rather than an indicator of non-sustainability. It is better to consider EFs as an attempt to assess the magnitude of global non-sustainability.
245. Less encompassing but more rigorously defined indicators, such as the “Carbon Footprint” (CF), would seem better suited to assess a particular dimension of sustainability, insofar as they are more clearly physical measures of a physical stock and of changes in that stock. (Again, they are better as measures of global sustainability, and there are complicated problems of assessment in general equilibrium models, e.g. in evaluating the carbon used to produce each of the inputs that go into each product. Marginal carbon utilizations may differ, for instance, from average utilization. Moreover, measuring sustainability requires an assessment of future changes in technology.)
246. One use of these indicators is to send strong messages concerning the over-utilization of the absorbing capacity of our planet. Because these metrics can be calculated at any level of disaggregation⁵⁹, they can be a powerful instrument for monitoring behaviours of individual actors.

c. Global perspectives

247. In a globalized world one needs to take a global perspective. This is obviously the case (and well recognized) for global warming, but it is also true for other depletable and renewable natural resources.
248. Taking a global perspective means that one cannot overlook the fact that some countries are more vulnerable than others to the degradation of the global environment. Regardless of where carbon gas is emitted, we know that climatic change has a strongly uneven distribution of consequences on different countries. Does worldwide sustainability mean the preservation of well-being for the *average* of the world population, or for those of its members that are more directly confronted by the adverse consequences of such an environmental threat? Whatever one's views on the long-term sustainability of the world's current path, for the Maldives, Bangladesh, and the low-lying Pacific Island states, the current path is not sustainable.

59. Subject to the qualifications made in the previous paragraph.

d. Other dimensions of sustainability

249. While environmentalists have performed a service in calling our attention to sustainability, the examples of non-sustainability resulting from excessive debt make it clear that these sustainability issues are far broader. Here too the present crisis helps us to understand other dimensions of sustainability, namely the human and social dimensions. A major crisis such as the current one or that of 1997-1998 can have long-term effects on the future welfare of some categories of populations, in particular in developing countries, as a result, for instance, of interrupted educations or malnutrition. And in turn, these consequences may jeopardize the sustainability of the social and the political system.
250. The discussion of sustainability serves to remind us that by looking at an economy today, we get a very incomplete picture of complex dynamic processes. Earlier, we emphasized the importance of the distribution of income. Again, we looked at a “snapshot”, a picture at a moment in time. But what matters for individuals is their lifetime income – especially if they can borrow to smooth out their income over time. Societies are, and should be, concerned with social mobility across generations.
251. In the development literature, emphasis is placed not just on economic and environmental sustainability, but also on political and social sustainability. We have no good metrics, for these other important aspects of society. Still, an important aspect of political sustainability is access to “voice” and social sustainability and equality of opportunity. Metrics on these variables can and should be developed. (See also the earlier discussion in Section XX)
252. An important warning should be sounded in closing. La Palisse would have said that a necessary condition for sustainability is that a catastrophe should not arise that prevents achieving sustainability. Climate change can precipitate the advent of such a catastrophe. Such a catastrophe would change dramatically the world in which we live, and therefore how we live—so much so that much of what we have said might seem irrelevant. We thus badly need indicators measuring how close we are to such a catastrophe. Our Report has been written by economists and social scientists who deal with numbers describing our economy and our society. We thus rely on scientists on climate change and other natural scientists to conceive these indicators of a potential catastrophe in our physical surroundings. We simply do not have the expertise to do this by ourselves.

THE WAY AHEAD: CONCLUDING REMARKS

253. The Members of the Commission believe that our work has just begun. We are convinced that what we measure and how we measure it makes a difference – for policy, for judgments about what works and what doesn't.
254. It should be obvious that no single number can summarize anything as complex and variegated as “society”. But, inevitably, certain numbers – in particular GDP – have taken center stage. Millions are spent forecasting what that number will be next month or next year.
255. The members of the Commission agreed that such a number may be misleading if it were applied to all purposes, and especially as a broader measure of societal performance. We have described a number of reforms that, depending on the purpose of the measure, would provide better metrics. Many of these reforms are attainable with relatively little investment in the collection and analysis of statistical data. And we are hopeful that our discussion of the strengths and limits of different metrics will lead to more care in the use of each, and better selection of the use of appropriate metrics for different purposes. For example, we have identified measures that are better adapted to the measurement of the well-being of citizens in a country in a globalized world than GDP; the evolution of median income is more telling about the representative agent than per-capita GDP.
256. Measurement of “present” economic performance also includes an assessment of “quality of life”. Quality of life includes the full range of factors that make life worth living, including those that are not traded in markets and not captured by monetary measures. While many of our measures are directed at ascertaining short-run movements in the level of market activity, the Commission considers that the time has come to make a clear move from measuring production to measuring welfare, to try to close the gap between our measures of economic performance and widespread perceptions of well-being. Building on existing research, it proposes new and credible measures for at least some aspect of quality of life.
257. Our societies have become more concerned about the environment, as we have realized that what we have been doing has placed our planet at risk. This realization reinforces the importance of sustainability, and heightens the imperative for the development of metrics that assess sustainability. While we have argued for the development of a broader measure of extended wealth to assess sustainability, we believe that such an approach should be supplemented with some physical measures.
258. We believe that a global debate around the issues and recommendations raised in the Report provides an important venue for a discussion of societal values, for what we, as a society, care about, and whether we are really striving for what is important
259. We hope that ours Report will provide the impetus not only for this broader discussion, but also for on-going research into the development of better metrics that will contribute to a better assessment of economic performance and social progress.