Conférences & Séminaires
03
December 2011

Measure for Measure

How Well Do We Measure Development?

Proceedings of the 8th AFD-EUDN Conference, 2010

Michel AGLIETTA
Sabina ALKIRE
François BOURGUIGNON
Andrew E. CLARK
Angus DEATON
Claudia SENIK

Measure for Measure
How Well Do We Measure Development?

Proceedings of the 8th AFD-EUDN Conference,
Paris, December 1st, 2010

AUTHORS

Michel AGLIETTA
Sabina ALKIRE
François BOURGUIGNON
Andrew E. CLARK
Angus DEATON
Claudia SENIK

DISCUSSANTS

Pramila KRISHNAN
Franck LECOCQ
Alemayehu Seyoum TAFFESSE

CONTACT

Robert PECCOUD
Research Department, AFD
peccoudr@afd.fr

Conférences et séminaires

AFD's Research Department organises a large number of seminars and conferences which provide a forum for exchanges of knowledge and experience between development aid actors: researchers, experts, political leaders, NGOs, companies... These events may address any of AFD's fields of operation. The Conférences et séminaires *collection aims to provide readers concerned by these issues with the main outcomes and lessons-learned from this research.*

Past publications in the *Conférences et séminaires* collection:

- *Implementing Large Scale Energy Programs in Existing Buildings in China*
- *Transitions décrétées, transitions vécues. Du global au local : approches méthodologiques, transversales et critiques*

All our publications are available at http://recherche.afd.fr

[Disclaimer]

The analyses and conclusions presented in this document are the responsibility of the authors and do not necessarily reflect the position of AFD or its partner institutions.

Publications Director:
Dov ZERAH
Editorial Director:
Robert PECCOUD

FSC
Designed and produced by Ferrari /Corporate – Tel.: 33 1 42 96 05 50 – J. Rouy / Coquelicot
Printed in France by: STIN

Conférences & Séminaires

Acknowledgements

This volume presents the proceedings of the conference "Measure for Measure: How Well do We Measure Development?", which was jointly organised by the *Agence Française de Développement* (AFD) and the European Development Research Network (EUDN) and held in Paris on 1st December 2010.

The French version of the eighth AFD/EUDN conference proceedings is published in the *Revue d'Économie du Développement* (2-3, septembre 2011).

For eight years now the Agence Française de Développement and the European Development Research Network (EUDN) have jointly organised a major conference on development. Hosted this year in the Ministry of the Economy, it enjoyed record attendance with over one thousand participants from some thirty countries in Africa, Asia and America.

With the years, this event now stands as one of Europe's leading fora for the development community. Each year, it raises the same challenge: bringing together a large and broad-based audience, including public institutions, businesses, NGOs, researchers, committed volunteers, etc., along with ground-breaking contributions from top experts. This event is, first and foremost, an opportunity for exchange. This has been the spirit of the previous conferences, which addressed major themes such as global inequality, international migration, the role of culture, managing natural resources under the pressure of population growth and the fragmentation of our globalised world.

In 2010, we saw fit to revisit a fundamental question: measuring development and, by extension, understanding the nature of development. All said and done, can we really know the meaning of "development" as long as there is no agreement on how to measure it?

We thank the members of the EUDN Network for their fruitful collaboration and the high quality of their contributions. We thank all participants whose presence largely contributed to the success of this eighth joint AFD/EUDN conference. We finally thank Louis Blazy, Philippe Cabin, Olivier Charnoz, Gill Gladstone, Sandrine Laborie, Véronique Sauvat and Laurence Wunderle, for their precious help in both the organisation of the conference and the editing of this volume.

Robert PECCOUD
Research Director
AFD

Conférences & Séminaires

Table of Contents

Acknowledgements 3

Introduction 7

by François Bourguignon

1 / Measuring Development: Different Data, Different Conclusions? 11

by Angus Deaton

Abstract 13
Introduction 14
1.1. More data, better data: benefits and challenges 18
1.2. The 2005 revision of the ICP and its consequences 24
1.3. Measuring hunger 37
1.4. From many to one: single indexes of development 46
References 49

2 / Multidimensional Poverty and its Discontents 53

by Sabina Alkire
Comment by Alemayehu Seyoum Taffesse

Abstract 55
Introduction 56
2.1. Multidimensional poverty 58
2.2. One particular application of M_0: the MPI 67
2.3. Research questions and debates 73
2.4. Concluding remarks 83
References 84
Comment: Multidimensional Poverty and its Discontents 91

3 / Will GDP Growth Increase Subjective Happiness in Developing Countries? 97

by Andrew E. Clark and Claudia Senik
Comment by Pramila Krishnan

Abstract 99
Introduction 100
3.1. The paradoxical relationship between growth and well-being 106
3.2. Explanations related to growth itself: channels and negative side-effects 130
3.3. Explanations related to the happiness function itself (human beings are social animals) 138
3.4. Conclusions and take-home messages: how can we use subjective variables in order to understand the GDP-happiness relationship? 156
Appendix 1. 159
Appendix 2. 161
References 163
Comment: Will GDP Growth Raise Well-Being in Developing Countries? 171
References 176

4 / Sustainable Growth: Do We Really Measure the Challenge? 177

by Michel Aglietta
Comment by Franck Lecocq

Abstract 179
Introduction 181
4.1. National income and social welfare: a perennial problem 184
4.2. The challenge of sustainability: from income to wealth accounts 188
4.3. Sustainability of public finances 195
4.4. Sustainable growth: measuring intangible capital 200
4.5. Sustainable growth: measuring natural capital 206
4.6. Discounting the future under radical uncertainty 217
References 222
Comment: "Sustainable Growth: Do We Really Measure the Challenge?" 226
Conclusion 229
References 230

Acronyms and Abbreviations 231

Authors' Biographies 235

Introduction

François Bourguignon, *Paris School of Economics*

The fact that economic growth is not necessarily synonymous with development and that GDP can grow in a country for some time without noticeable improvement in the areas of poverty, health or education has long been known. Yet development continues to be measured primarily by GDP growth and public attention continues to focus on this statistic alone. Likewise, the fact that populations increasingly participate in market activities is taken as a sign of progress, even though it may come at the cost of shrinking non-market activities. The effectiveness of development aid also continues to be evaluated on the basis of GDP growth, although its impact on health or education may be more important in the medium-run.

There are countless examples of the inadequacy of the GDP indicator to monitor development. The key point is that GDP is no more than a measure of market activity within a country. In that sense, it is an essential economic concept and a most useful indicator. But it does not measure development in all its dimensions: material well-being, availability of public goods, quality of life, social inclusion, access to justice, collective decision-making.

But then, how should we measure and evaluate development? Should we define a multiplicity of indicators that describe its various aspects, ranging from access to drinking water to the quality of the judiciary system, from infant mortality to the democratic features of society, from the calorie intake of the average citizen to reported individual satisfaction or happiness? Today, international comparisons that use this type of indicator abound. What they mean, however, remains ambiguous These international comparisons sometimes attract the public's attention but they are often more like statistical games with no strong or obvious implications in terms of development.

Should we then try to integrate all the dimensions of development and define an aggregate indicator that would include all of them? This "super-GDP" would go far beyond an indicator such as the Human Development Index (HDI) used by the UNDP and presently being extended to account for inequality. Although attractive, this approach is not necessarily the soundest. The various dimensions of development are simply not comparable, and it is impossible to imagine them being interchangeable as would be implicitly the case within an aggregate index.

Such are the questions that were addressed by top international experts in this joint AFD/EUDN conference on "Measuring Development". Each of them, in their own field, will focus on approaches that can be taken as alternatives or complementary to traditional statistical measures not only to quantify economic and social progress, but also to design and monitor better development strategies.

The issues to be covered notably included the increasing use made of various types of household surveys and their limits, whether the focus is on economic well-being, health or poverty when defined as a truly multi-dimensional concept. Attention was also drawn to the new indicators that summarise the various aspects of development through the average individual subjective satisfaction of a population and how this may compare with more traditional objective measures. On a different topic, special focus was given to the issue of "development sustainability", in particular through the concept of "net effective savings", which takes into account the consumption or depreciation of various stocks of natural, material, human or social resources.

The specific topics covered by the various papers and discussions in this volume are:

International pricing, national accounts, and household surveys: how can these data inform us on standards of living, poverty and inequality?

Today, we have more and better quality measures of economic development than ever before. Household surveys are constantly growing in number and availability, providing us not only with data on household income and expenditure, but also direct measures on health, particularly on anthropometrics, infant mortality, and self-reported material and emotional well-being. The latest round of the International Comparison Project (ICP) collected prices of comparable goods in 146 countries, many of which had not yet come under study. These data bring new ideas and new insights on national and individual economic development. Yet, there are persistent problems of interpretation and consistency between these different types of data. Why is world poverty not retreating as quickly as one might expect in view of global economic growth? Why are Indians consuming fewer and fewer calories when their nutritional status is so low and their incomes rising rapidly? Why does economic growth not always tie up with an improved feeling of well-being in a population? And how should we interpret the pronounced increases in poverty and inequality estimates shown by the ICP.

The multidimensionality of poverty

The multidimensionality of poverty is very inadequately accounted for in conventional statistics. The indicator most often used is "monetary poverty", which is generally defined as monetary income (or consumption) below an arbitrarily set threshold. However, other kinds of deprivation, such as no access to basic infrastructure, health care, education or public decision-making are equally important – and do not neatly correlate with income.

So far in the research on poverty, the spotlight has been on "monetary poverty", while data on access to safe drinking water, infant mortality, illiteracy or political participation have been treated separately. Recent studies have tried to remedy this approach by integrating the different dimensions of poverty. In line with Sen's "capability" approach, they explore different types of deprivation that reflect the various facets of poverty. They also pay particular attention to areas where these different dimensions may overlap when one is assessing the severity of poverty.

The 2010 UNDP report on human development illustrates this approach. It shows more specifically that multidimensional measures of poverty based on a small number of deprivations lead to international rankings that may differ substantially from rankings based solely on monetary poverty – and of course on GDP.

Measuring development through individual subjective satisfaction

Defining a single indicator that aggregates the different dimensions of development presupposes that "weighting" indices have been fixed for calculating a weighted sum of indicators corresponding to the various dimensions. However, it is almost impossible to reach a consensus on what coefficients should be used for this weighting. On the other hand, the subjective satisfaction indicators can be taken as producing exactly this type of aggregation at the level of each individual in society. Average individual satisfaction thus implicitly provides a solution to the problem of aggregating the different dimensions of development.

Considerable work has recently been done on all these data on satisfaction, sometimes also dubbed "happiness" data. The question is knowing to what extent these data can actually be used as development measures, and serve cross-country or intertemporal comparisons. More importantly, the pivotal question is knowing how they correlate with traditional development indicators.

According to the Easterlin paradox, an unexpected rise in a country's income *per capita* improves subjective satisfaction, but only for a time. In the near-to-mid run, a correlation between income and satisfaction does indeed exist. But in the long run, people become used to their new level of income and their satisfaction ends up no higher than it was before their income rise. It would be useful to know whether this paradox also applies to other aspects of development and well-being and if it is present in all cross-country comparisons.

Sustainable growth: do we really measure the challenge?

If the project of sustainable growth is to be taken seriously, this poses hefty challenges for economic science, for our accounting systems and for economic policy choices. After decades of capitalist accumulation underpinned by the unbridled exploitation of natural resources, it is now recognised that these may not be infinite, that the stability of our ecological systems, and by the same count the well-being of future generations, are far from secure. These goods therefore take on a value. But how can this value be measured seeing that there is no historical precedent, no corresponding market, no commonly agreed norms to set priorities? What discount rates should apply? Can environmental goods be substituted for other goods? The questions are many and the debate continues to swell. At the heart of the sustainable growth revolution is the challenge of mapping and measuring hitherto largely unexplored terrain. The definitions of capital, savings, investment and welfare need adjusting to reflect the importance of intangible capital, natural capital and the uncertainty that threatens the stability of the Earth's climate system. This means no less than an overhaul of our entire definition of wealth and thus our changing national accounts system and statistical machinery, in order to accommodate these dimensions.

The reader might think that these diverse papers raise more questions than answers. In a domain whose complexity is becoming more apparent with each passing day, this is more or less inevitable. However, the overall approach is still useful. Limiting ourselves to a few simple development indicators, as we often tend to do, simply because they are relatively easy to observe, boils down to looking for our keys under a lamppost because there is more light there. Thinking on development over recent decades has partly involved identifying these different dimensions and shedding more light on how they interact. "Measuring development" now needs to take stock of the progress made with these dimensions, be it economic activity, health, education or the environment, to understand their origin and implications and, finally, to make some kind of synthesis. The papers that follow highlight the progress achieved in this direction and also point out some key shortcomings for which our metrics must try and find an answer.

Conférences & Séminaires

1. Measuring Development: Different Data, Different Conclusions?

1. Measuring Development: Different Data, Different Conclusions? [1]

Angus Deaton, Princeton University

Abstract

We now have more and better measures of economic development than ever before. The number and availability of household surveys have been improving over time. These surveys provide data, not only on household incomes and expenditures, but also on direct measures of health, particularly on anthropometrics, on infant and child mortality, as well as on self-reported measures of well-being and emotional experience. It is possible, for the first time, to compile global maps of multiple components of human welfare.The latest round of the International Comparison Project (ICP) has collected prices of comparable items in 146 countries, many of which have not been previously surveyed. These new data have brought many new insights and new discoveries about economic development of both nations and of individuals. Yet there are also problems of interpretation and consistency between the different types of data. Why does world poverty not fall as fast as might be expected given the amount of growth in the world? Why are Indians consuming fewer and fewer calories when their nutritional status is so poor, and their incomes are rapidly rising? Why is economic growth not always associated with improvements in self-reported well-being? And how should we interpret the marked increases in estimates of global poverty and global inequality that came with the latest data from the ICP? This paper reviews these puzzles and questions and identifies key questions that need to be resolved.

[1] I am grateful to Olivier Charnoz, Eric Jourcin, Robert Peccoud and Cecile Valadier for comments on a previous draft.

Introduction

More data, better data and a broader perspective

The measurement of economic development has changed beyond recognition in the last twenty years; old measures have been improved, made available for many more countries and with greater frequency, and many new measures exist. At the same time, the concept of economic development moved on from an exclusive focus on growth in real incomes – and a view of poverty and deprivation as a lack of real income – towards the inclusion of other dimensions of human welfare, among which health has received the most attention. This conceptual change owes a great deal to the work of Amartya Sen who in *Development as Freedom* (1999), emphasizes that these multiple dimensions are not only *components* of welfare, but also interact as causes of development and deprivation. The change was recently given further impetus by the work of the Commission on the Measurement of Economic and Social Progress (Stiglitz, Sen and Fitoussi, 2009), which recommended the systematic incorporation into official statistics of broader conceptions of welfare, supported by many new measures. The new and better data are the basis for an explosion of work, not only in measurement, but also in the investigation and understanding of mechanisms, particularly those linking income and health. Yet new information often poses challenges, in understanding why it contradicts previous perceptions, or why it appears to undermine what are seen as well-established regularities.

Three topics: prices, poverty and inequality; hunger; and health

In this paper, I review recent developments in measurement and identify several outstanding puzzles and questions. I focus on three specific areas. First is the most recent revision of the International Comparison Program (ICP), benchmarked on the year 2005, and published early in 2008 (World Bank, 2008). These new numbers – price indexes based on millions of prices from 146 countries – changed our view of the world, moving poor countries further away from rich countries and so expanding measured world inequality. They were also accompanied by a major upward revision of the number of poor people in the world. My second topic is the measurement of global hunger, a topic that attracted a great deal of attention when the Food and Agriculture Organization of the United Nations estimated that the food price spike in 2008 and the financial crisis of that year has led to an increase of nearly 200 million people in hunger (FAO, 2009). I discuss the origins of such numbers, question their relevance, and present some alternative, new calculations.

Third and finally, I turn to the question of how to think about health and income together. I argue that when we are concerned with *measurement*, multidimensional measures are what are required, and that these, as in Alkire and Santos (2010), need to be calculated from surveys that collect multiple measures for each respondent. Measures that are computed from national averages ignore one of the most important aspects of poverty and deprivation, that deprivations in different

dimensions are positively correlated: people with low income are typically also people with poor health, poor access to education, and less than full participation in political and civil life. However, when we go beyond measurement and try to understand the causes of poverty, it is essential to keep the different measures distinct, and to resist the temptation, in spite of the correlations, to use one dimension as a proxy for another. There are many instances where health and income do not go together, and to take income as an indicator of health, or improvements in health as an indicator of economic growth misses the reasons why they are different. Healthcare policy and innovation in healthcare provision are both capable of improving health in the absence of economic growth (or of failing to do so in its presence) so that it is often the divergence between the measures that identifies the importance of policy and of innovation.

The main arguments: a guide to the paper

The paper covers a lot of ground, and the arguments are sometimes detailed. So it is useful to anticipate the main conclusions and link them to the sections where they are discussed in detail.

• More than at any time in history, we have a wealth of data from most of the countries around the world. Although there are gaps, we now have an unprecedented collection of data on prices, incomes, health, and well-being. In many, although not all, cases, these data are collected on a comparable basis so that there are new opportunities for the global mapping of human welfare (section 1.1.1.). Better data also raise a number of puzzles and contradictions (section 1.1.2.).

Prices, poverty and inequality

• The price data from the latest round of the International Comparison Project are better and more comprehensive than those from any previous round. However, the high quality of these data has also clarified a number of remaining issues (section 1.2.). The ICP is used to convert national income estimates to "real" comparable units, so the quality of those numbers are only as good as the underlying national accounts which, in many cases, are weak. Improving national accounts should be prioritized by the international community.

• Improvements in the ICP have clarified the conceptual difficulties of making real income comparisons between widely different economies. Real income comparisons between even major countries, such as the US and India, or Britain and China, are subject to much larger margins of uncertainty than are commonly recognized (section 1.2.1.). When people in different countries have different patterns of consumption, there is no non-arbitrary way of calculating cost-of-living index numbers with which to compare them.

• Global poverty estimates use a common international poverty line that is defined as the average of poor country poverty lines. Given revisions of the PPP exchange rates with each ICP, and given revisions of the countries in the average, the global poverty counts are subject to large revisions, in the most recent case, an upward revision of half a billion people that has made global poverty more Asian, and less African. Such moving targets undermine any serious program for international poverty reduction (section 1.2.2.). Given the current procedure for defining the global line, such revisions cannot be avoided.

• The PPPs are currently revised only with each new round of the ICP. A more continuous process of revision, depending on exactly how it is done, could modify or even reverse the rate of global poverty decline, and could convert decreasing global inequality to increasing global inequality. Once again, there is much more uncertainty than is commonly recognized (section 1.2.3.).

• Poverty counts and inequality measures are undermined by major discrepancies between national accounts and household surveys in many countries, not only in levels, but also in rates of growth. Poverty, as measured from household surveys – as done by the World Bank – declines less rapidly than would appear to be warranted by the amount of economic growth in the world. Reconciling national accounts and household surveys should also be an international priority, though there are a number of political and statistical obstacles (section 1.1.2.).

Hunger

• Given the uncertainties associated with comparisons of real income, poverty counts, or global measures of inequality, as well as for substantive reasons, there is much to be said for paying attention to other measures of welfare. One such is whether or not people are well-nourished. There are two classes of hunger measures, undernutrition – people not having enough to eat – and malnutrition – people's bodies showing the signs of inadequate nutrition, for example by being too thin or too short (section 1.3.)

• The hunger measures produced by the FAO are undernutrition measures, which calculate – or forecast – whether incomes and food prices will allow people to buy what they need. The "flash" numbers are entirely forecasts, but I develop independent evidence (from Gallup's World Poll, which asks people if they have enough money to buy food) that confirms at least some of the spike in undernutrition in 2009 (sections 1.3.1. and 1.3.4.).

• Direct measures of malnutrition – based on the measurement of heights and weights – are both conceptually and substantively different from the measures of undernutrition. Because these data come only with a lag, we do not have data for the most recent years, but the geographical pattern of malnutrition is very different from the geographical (largely income-related) pattern of undernutrition. On average, malnutrition is much worse in South Asia than in Africa, in spite of higher levels of income and lower levels of undernutrition; the reasons for this are not well understood. Measures of deprivation that include measures of malnutrition in addition to income poverty further shift the prevalence of poverty from Africa to South Asia (section 1.3.2.).

• The complexity of the relationship between income and nutrition is illustrated by current trends in India, where rapid economic growth, together with poverty reduction, have been accompanied by *declines* in *per capita* calorie consumption, in spite of some of the world's highest levels of malnutrition (section 1.3.3.).

Health and health & income (section 1.4.)

• There is legitimate demand for indexes that combine health and income measures into a single index. Standard methods of combining means are much inferior to methods, such as the new multidimensional indexes, that aggregate at the individual or household level, though the latter has more severe data requirements.

• For understanding the process of development, it is important not to conflate health and income (or other components of well-being) because they do not always move together, often precisely because of the government policies whose effects we need to understand.

Self reported well-being (section 1.1.2.)

• Self-reported well-being (SWB) measures have recently received a great deal of attention. While they are often useful, they need to be treated with skepticism, if only because adaptation can make them unreliable guides to objective deprivation.

• Contrary to much of the literature, it is important to distinguish different measures of SWB. In particular, life evaluation measures behave differently than emotional measures. It is likely that life evaluation is less subject to adaptation than are measures of emotional well-being, and thus arguably more suitable as a measure of development.

• The Easterlin paradox, that economic growth is not accompanied by improvements in well-being is still alive, if under increasing attack. Until it is resolved, it is hard to recommend SWB measures as a gauge of economic development.

Political economy of global measures

• Global measures of development – poverty, inequality, hunger, or price levels – operate in an entirely different political environment than do domestic measures. The latter, for example domestic consumer price indexes, feed into domestic policymaking, and are typically subject to oversight procedures that constrain both the statisticians who produce the data and the politicians and policymakers who use them. The international agencies who produce global statistics are subject to no such oversight, and so are not protected against even ill-founded suspicion that they manipulate the numbers in their own interests. The World Bank's upward revision of 500 million people in poverty is of a magnitude that is hard to imagine in any important domestic statistic, and the lack of any major reaction from the international community suggests that global measures play little or no role in international policymaking; if so, their significance is unclear.

• If the international development community believes that global measures of development are important, it should consider better monitoring and oversight of the production of the most important measures.

1.1. More data, better data: benefits and challenges

1.1.1. Surveying the data landscape

In his famous 1955 paper on income inequality, Simon Kuznets had data for only six countries: Britain, Ceylon, Germany, India, Puerto Rico and United States. In the first ever paper on counting global poverty, Ahluwalia, Carter, and Chenery (1979) had sufficient distributional data to calculate poverty rates for 36 developing countries. They also used data on purchasing power exchange rates from the first two phases of the ICP published in Kravis, Heston and Summers (1978) to calculate a global poverty line. Kravis *et al.* used ICP data that were benchmarked (meaning there were actual price data) for 16 countries, and extrapolated their results to a total of more than 100 countries. By contrast, for the latest round of the ICP, the World Bank (2008) collected prices for 146 countries, and Chen and Ravallion's (2010) most recent counts of global poverty use almost 700 household surveys from developing countries, many of which, like the Bank's Living Standard Measurement Surveys, collect data not just on income and consumption, but on health, education, child mortality, anthropometrics, calorie intake, and a host of other topics.

There have also been major advances in the collection and availability of data on health, although major gaps remain. The system of Demographic and Health Surveys (DHS) has evolved from what was originally an almost exclusive focus on reproductive health. The contemporary DHS collects data not only on reproductive histories – which are the basic material for estimates of infant and child mortality in countries without complete vital registration systems (the majority of poor countries) – but also collects weights and heights, at first for children, then for adult women, and most recently for adult men. Such data have been collected piecemeal in some countries; for example, India has a national nutritional monitoring bureau that covers only part of the country in some years, and which has used different standards in different surveys. But the DHS system uses comparable questionnaires in different countries. This may not always be ideal for the country, but it is a boon for researchers who are beginning to paint something like a complete picture of nutritional status around the world measured, not by food intake, but by physical outcomes. The DHS also collects information on the ownership of a range of durable goods which, following Filmer and Pritchett (1985), has become a widely used measure of economic status in the absence of questions on wealth, income, or expenditure (for which it is a far from perfect substitute). I should also note the role of the DHSs in testing for HIV-status, a program that caused a major reassessment of global prevalence and its distribution across countries.

At the aggregate level, the WHO collates and makes available national data on mortality rates by age, sex, and cause of death. These are of most use for the richer countries of the world

which have the complete vital registration systems from which the WHO mortality database is assembled, though there are good data for a number of middle-income countries, for example in Latin America, and one or two poor countries, such as Sri Lanka, which have exceptionally complete data. India and China do not have complete vital registration systems, but have other methods of compiling national estimates of mortality by age and sex, though they lack the detail that is available for the OECD countries. The absence of adequate adult mortality data for most poor countries, including almost all of sub-Saharan Africa, remains the most glaring deficiency in the system of global health statistics. Christopher Murray and his colleagues at the Institute for Health Metrics have recently pieced together all of the fragmentary data that is available to provide a set of new estimates of child and adult mortality by cause of death (Rajaratnam *et al.*, 2010a, b). These rely heavily on imputations, for example, from small areas with good data to large areas with none, and while these numbers are almost certainly the best that can be done, they should be treated with caution and should not disguise the underlying absence of hard numbers. In particular, in most of the places where adult mortality is highest, we do not have the kind of data that is required to monitor and evaluate local and international health interventions.

Another rapidly expanding area is the measurement of self-reported well-being (SWB), or what is often called "happiness", though this designation can be seriously misleading. The World Values Surveys, beginning in 1990, have asked a range of life-satisfaction questions; in the first wave, these surveys were not nationally representative in the relatively few poor countries included (deliberately so) but this has been progressively rectified in waves 2 through 4, the last of which was collected in 2005; a 2010/11 wave is currently in the field. While this will provide a 30-year series for many countries, the analysis of change in poor countries is dangerous because of the changes in selection. There are also systems of "Barometer" surveys for Europe, Latin America, Asia, and some countries in Africa, some of which collect SWB data.

An important new entrant into this area is the Gallup World Poll, whose ambitious aim is to provide ongoing monitoring of all of the people in the world. Begun in 2006, it has so far collected data in more than 150 countries, although not every country is covered in every year. The World Poll is distinguished by the fact that the identical core questionnaire is given to all respondents in all countries; while this limits the range of topics, it provides an unusual degree of international comparability. The questionnaire is administered by phone in rich countries, and face to face in poor countries, and the questions have been tested and tailored to avoid mode bias; the samples are typically 1,000 respondents, although sometimes larger, and except in a few cases where regions of countries are inaccessible, are nationally representative. The World Poll asks a number of questions about self-assessed economic status, one of which on not having enough money to buy food, I shall use in section 1.3. below. It is also unusual in having an array of different questions about self-reported well-being, so that it is possible to distinguish between hedonic well-being (happiness, enjoyment, sadness, stress, etc., as experienced yesterday) and life evaluation, which asks people to think more broadly about how their lives are going. The Gallup

Organization undertakes the World Poll as a commercial venture (why did none of the international agencies collect this kind of data?), which has the disadvantage that the data are not publicly available.

1.1.2. Puzzles and challenges

The new round of the ICP has raised many issues, if only because the picture of the world that it paints is so different from the picture that was previously familiar. It not only gives us a new and much more unequal world, but presents us with the problem of how to link it with the old world. Can we simply accept the new shape of the world, together with the old rates of growth, and discard the old shape altogether, which would be appropriate if the new data simply corrects errors in the previous data, and as is done in the World Bank's widely used World Development Indicators. Or was there some truth to the old estimates, so that we need to change our views of growth too? I will deal with some of these questions in section 1.2.

Surveys and national accounts

The expansion in the number of household surveys has also highlighted an issue that has been long known in individual countries, including India and the United States, but which appears to be of much wider applicability (see Deaton [2005] for a full account on which the following summary is based). It turns out that the surveys are generally inconsistent with the national accounts, both in the structure of expenditures over groups of goods and services, but also in their estimates of the rate of growth of *per capita* consumption over time. The former matters (among other things) for the construction of index numbers, such as purchasing power parity exchange rates, while the latter matters for the measurement of poverty. On the almost certainly correct assumption that the errors are not only in the surveys, the discrepancy also casts doubt on the measurement of aggregate consumption and GDP. For example, in both India and the United States, *per capita* consumption estimated from the household surveys rises one percent a year more slowly than does *per capita* consumption measured in the national accounts. Some, but not all of the discrepancy can be attributed to differences in coverage and in definition; there are many imputed items in the national accounts – imputed rent for housing and financial intermediation indirectly measured are two of the most important – none of which show up in the surveys. It is almost certainly true that the surveys are missing progressively more expenditures over time, perhaps because the responsibility for spending is more widespread over household members than it used to be, so that the "single knowledgeable respondent" mode of interviewing misses more and more. In the US, there are many cross-checks on most aggregate consumption items in the national accounts, so that the burden of proof tends to fall on the surveys. But the quality of national accounts is much lower in many poor countries, with many numbers little more than guesses, so there is no such presumption internationally. One study in India by official statisticians, Kulsheshtra and Kar (2005), looked at discrepancies in food categories, and while there was plenty of blame to go round, the surveys were more often judged to be correct. In Deaton (2005), I argue that there are reasons to suppose that national income accounting procedures tend to overestimate growth rates when growth rates rise, for example by double-counting intermediates by using short-cuts that were designed to work at lower levels of

income. But it is difficult to persuade the governments of rapidly growing countries to risk downgrading their own success by digging too deeply into their national accounting practices.

That survey means grow more slowly than the corresponding means in the national accounts also makes mischief with the measurement of poverty. In early poverty measures, such as the Ahluwalia *et al.* (1976) or the government of India's own procedures, and in historical reconstructions of global poverty, most notably Bourguignon and Morrisson (2002), poverty was estimated from the national accounts data, supplemented by distributional data from the surveys. For example, the combination of an assumed lognormal distribution whose variance is estimated from a survey and mean *per capita* consumption from the national accounts identifies the position of the distribution, and yields the fraction of the population below any given poverty line. Most contemporary poverty estimates, including the World Bank estimates, work directly from the surveys, and calculate the headcount ratio from the actual empirical distribution, without reference to the national accounts. When survey means are growing more slowly than the means in the national accounts, the "old" procedure will show more rapid poverty decline than the "new" procedure. Without an as yet unattained understanding of the differences between the two sources, we have no way of deciding which rate of poverty decline is correct. Several authors, most notably Bhalla (1997) and Sala-i-Martin (1998) use the "old" procedures, and (unsurprisingly) claim that the World Bank estimates, which use the "new" procedures, are understating the rate of poverty decline. In a more recent paper, Pinkovskiy and Sala-i-Martin (2009) use a variant of the same method, and find that their estimates of rapid poverty decline are robust to a wide range of variations in their assumptions, except the crucial one of replacing national accounts means by survey means, which they do not consider. None of these papers offer a rationale for believing that national accounts are correct and surveys wrong, nor do they explain what assumptions are required to justify discarding the survey mean while accepting survey measures of dispersion; one possible account is given in Deaton (2005), but it requires special assumptions whose validity is far from obvious. So it seems unlikely that these optimistic calculations are correct, though it is also most likely true that the Bank poverty estimates, which are based entirely on surveys, understate the rate of decline of income poverty.

Hunger, nutrition and mortality

Another set of contradictions arises in the measurement of hunger and nutrition. One important distinction is between undernutrition – which refers to people not having enough to eat – and malnutrition – which refers to people being physically underdeveloped, by being too thin or too short (or both) or, in extreme cases, showing clinical signs of malnutrition, such as edema, marasmus or kwashiorkor. Undernutrition is measured either by collecting food consumption data and converting them into calories, protein, fat, and micronutrients, or through specialized nutritional surveys that directly monitor individual intakes of food. Malnutrition is measured by anthropometric measurement of height and weight for adults and children; these measures are usually included in dedicated nutritional surveys, but rarely in household expenditure surveys. Data from many countries are now avail-

able through the DHS system, as well as through UNICEF's Multiple Indicator Cluster Surveys (MICS), and the WHO's World Health Surveys (WHS). In a straightforward world, food consumption would rise with income, calories and other nutrients with food consumption, and both undernutrition and malnutrition would fall along with income growth. Across countries, undernutrition and malnutrition would be lower in richer countries than in poorer countries. Unfortunately, the world is a good deal more complicated, and none of these propositions is generally correct. I elaborate and discuss some of the possible reasons in section 1.3. below.

Rates of infant and child mortality are important indicators of deprivation and, in the absence of adult mortality data, are used to estimate variations in life expectancy at birth, albeit with some allowance for mortality from HIV/AIDS. Mortality rates and life expectancy are closely related to income across countries, but once again there are puzzles: child mortality is much higher in sub-Saharan Africa than in South Asia, in line with income differences, while malnutrition is lower in Africa. The rate of economic growth strongly predicts proportional changes in child mortality, but not absolute changes, essentially because economic growth is higher and child mortality lower in the richer countries. Increases in life expectancy in rich countries have recently been driven by decreases in mortality among middle aged and elderly adults, while increase in life expectancy in poor countries – other than those affected by HIV/AIDS – have been largely driven by decreases in infant and child mortality. These patterns have implications for how we think about and measure overall well-being, as well as for thinking about policy. I turn to these questions in section 1.4.

Self-reported well-being

I close this section with some remarks on the measurement of "happiness", or better, the measurement of self-reported well-being (SWB). The topic is dealt with elsewhere in this set of papers, so I can be brief. Routine measurement of SWB is recommended in Stiglitz, Sen, and Fitoussi (2009), a third of which is devoted to the topic. At its most ambitious, "happiness" responses are treated as definitive measures of human well-being, and the maximization of total measured happiness becomes the only criterion for public policy, views that are endorsed – with only minimal qualification – by Layard (2005). But most writers in the field have expressed greater skepticism. Indeed, there are good grounds for not accepting self-reported well-being as definitive at all, grounds that are perhaps particularly relevant in the context of assessing poverty and deprivation. Sen writes: "*a person who has had a life of misfortune, with very little opportunities, and rather little hope, may be more easily reconciled to deprivations than those raised in more fortunate and affluent circumstances. The metric of happiness may, therefore, distort the extent of deprivation, in a specific and biased way. The hopeless beggar, the precarious landless laborer, the dominated housewife, the hardened unemployed or the over-exhausted coolie may all take pleasures in small mercies, and manage to suppress intense suffering for the necessity of continuing survival, but it would be ethically deeply mistaken to attach a correspondingly small value to the loss of their well-being because of this survival strategy*" (1987, pp.45-6).

This ethical mistake can be avoided by following a capabilities approach, by which we measure aspects of capabilities – income, life

expectancy, malnutrition – without necessarily expecting to be able to combine them into a complete ordering (Sen, 2009).

Sen's concerns must be taken seriously, but whether or not SWB measures have the sort of bias identified by Sen is an empirical matter, at least in part. Nothing rules out the possibility that some SWB measures are good indicators of capabilities, and even if they cannot serve as overall indicators, they are certainly important measures in their own right: it is surely better to be happy than sad, to be carefree than to be worried, and to perceive one's life as going well rather than badly. On this empirical evidence, the jury is still out. In particular, there is no complete resolution of the Easterlin (1974) paradox that at least some measures of SWB have not increased with economic growth, although Stevenson and Wolfers (2008) have made some progress in that direction. If economic growth brings no increase in SWB, most economists still tend to believe that this reveals the deficiencies of SWB measures, and not follow Easterlin and Layard into the belief that economic growth does not improve the human lot. It turns out that it is important not to treat all SWB measures as the same because they correspond to different aspects of well-being. In particular, measures of momentary affect (or affect yesterday) capture current hedonic well-being – the experiences that make up the emotional texture of life – while life evaluation measures capture, not people's current feelings, but how they think about their lives, the distinction between experiencing life and thinking about it (Kahneman and Riis, 2005). Across countries, the Cantril life evaluation measure (a scale of 0 to 10 from the worst possible life to the best possible life) is astonishingly well predicted by (the logarithm of) *per capita* GDP, both among individuals and national averages (Deaton, 2008). Within the contemporary United States, hedonic experience responds to household income, but satiates at an income level of around $75,000, whereas life evaluation continues to rise with income (Kahneman and Deaton, 2010). There is at least the possibility here that life evaluation measures do respond to economic growth over time, which would help resolve the Easterlin paradox, although we do not yet have long enough time series of the Cantril measure to know.

1.2. The 2005 revision of the ICP and its consequences

1.2.1. Background

The rounds of the International Comparison Project are like successive Olympic Games. Like the Olympics, they do not happen every year, and in the first modern games only a few countries sent competitors. There were only a few events, and standards of competition were relatively low. The participants were amateurs with day jobs, and while they were great natural athletes, they did not take their training very seriously. Yet the first modern Olympics was a watershed, which eventually grew into the record-breaking, professional event that it is today, in which almost all of the nations of the world come together into a truly global competition. The ICP began in the late 1960s and early 1970s with Alan Heston, Irving Kravis, and Robert Summers from the University of Pennsylvania, and Zoltan Kennessy from the United Nations. The first round in 1967 had only six countries with four more added in 1970, and prices were collected for only a small range of goods and services. Since then, each round had become bigger and better (and more expensive), with more countries represented, with the involvement of more and more professional statisticians and economists, and with lots of preparatory training in the form of expert workshops, theoretical papers, and figuring out how to deal with problems that could not be solved in the previous round. The 2005 round was by far the most professional, the biggest, the most thoroughly researched, and the most international – with 146 countries.

ICP 2005 incorporated many improvements over the previous round in 1993, and perhaps the simplest summary is that the statistical procedures were so much better that the new estimates of PPPs are not really updates of the old, but a whole new set of incomparable numbers. In 1993, many countries had their PPPs imputed, because no price data were collected for them; these absentees included both India and China. The definitions of commodities and services were much more carefully specified in ICP 2005. And perhaps most importantly, the regional structure of the ICP was complemented by a strong global office, run by the World Bank, which developed and implemented a coherent plan for transforming a system of regional PPPs into a global set of estimates. The 1993 round was not centrally coordinated or controlled and, in the face of underfunding at the center, became a set of regional exercises, carried out at different times, each of which collected data and calculated regional PPPs. A UN report in 1997, under the chairmanship of Jacob Ryten, concluded that the estimates from ICP 1993 were not credible and concluded, with faint praise, that "*the ICP is a programme worth keeping but that its current condition, if little is done about it in terms of credibility, quality of output, and survival prospects, is poor*".

The linking of the regions in the ICP 2005 is not without its problems. The most serious of these are not failings of the ICP itself, but reflect conceptual differences in making com-

parisons between countries whose patterns of consumption and relative prices are radically different from one another. It is one thing to make PPP comparisons of France and Germany, or of Kenya and Tanzania, but we are on altogether more difficult ground when we come to compare Canada with Cameroon, Japan with Senegal, or Bolivia with Tajikistan. For example in Deaton (2010), I discuss the case of Cameroon and the U.K., whose bilateral price comparison is a component of the broad regional PPPs that link the regions. Air travel is very expensive in Cameroon, but its share in consumption is very small, so we might expect the high price to play little role in the bilateral comparison. But the price indexes that compare Britain and Cameroon use weights that are averages of the weights in the two countries, so the high price in Cameroon attracts half of the large British weight on air transport, and plays a significant part in the overall PPP. The relevance of such comparisons for the average citizen of Cameroon can be doubted, let alone for someone living at the global poverty line. More generally, the goods that are chosen for comparing across countries should be both truly comparable and widely consumed in both countries, criteria that are often in conflict. These and other outstanding issues for the ICP are discussed in Deaton and Heston (2010).

Finally, I note an important issue that is sometimes misunderstood. The ICP collects data on prices; it does not collect data on the national accounts of the participating countries. Although the ICP may sometimes lead to technical improvements in national accounts, the ICP's price indexes depend on weights from the national accounts, and its estimates of consumption or GDP at international prices come from deflating country estimates in local currency by the ICP's PPP exchange rates. A broader ICP might one day collect information on quantities as well as prices, but it does not do so today, nor did it do so in the past. The ICP quantity comparisons are only as good as the national accounts that go into them, over which the ICP has no direct control.

As elaborated below, the consumption PPPs play an important role in the calculation of the World Bank's global poverty counts. An often-heard criticism is that the weights for these PPPs are the aggregate weights from the national accounts, which do not reflect the consumption patterns of the poor. While that criticism is correct in principle, the reworking of the weights for PPPs in Deaton and Dupriez (2011) shows very little difference. While it is true that the weights for the poor are different from the aggregate weights, the difference does not vary very much across countries, leaving the price indexes largely unchanged. A larger difference comes from replacing the weights from the national accounts by weights from household surveys, taking us back to the contradiction between them.

The rest of this section is devoted to the consequences of the revision of the ICP for the measurement of global poverty and global inequality.

1.2.2. Measuring global poverty

History of global poverty measurement

The first calculations of global income poverty in anything like modern form are contained in Ahluwalia, Carter, and Chenery (1979). They use purchasing power parity exchange rates from Phases I and II of the ICP, centered on 1970, and updated to 1973, and published in Kravis, Heston and Summers (1978). These are used to convert an Indian poverty line into 1970

international dollars. The line is $200 *per capita* per annum, which is described as being about the 45th percentile of Indian GDP *per capita*, chosen as the middle of the range of 40–50 percent, which were the then current estimates of the headcount ratio in India. Although the calculations are not described in any detail, it appears as if the distribution of *per capita* expenditure from household surveys was applied to the total of GDP *per capita*. Note that the $200 line is *high* relative to subsequent global lines. The World Bank does not currently publish estimates of Indian GDP in 1975 in 2005 constant international dollars, but we can piece together growth rates from the *World Development Indicators* and from the Penn World Table (PWT) 6.2, which suggest that the 1975 figure in 2005 international dollars was around $764. Ahluwalia, Carter, and Chenery's poverty line of $200 is two-thirds of *per capita* GDP in 1975, so that their poverty line is $509 in 2005 international dollars, or about $1.40 a day. One reason for the line being so high is presumably that it is anchored in GDP *per capita* from the National Accounts, rather than *per capita* consumption expenditure from the Indian surveys, which is a much lower number.

The World Development Report (WDR) of 1990 is the source for the original $1-a-day line. The calculations move on from the 1970 round of the ICP to the 1985 round, the results of which were available in version 5 of the Penn World Table and described in Summers and Heston (1991). The report works with two lines, $275 and $370 per person per year ($0.75 and $1.01 per day) in 1985 international dollars. The text says that "*this range was chosen to span the poverty lines estimated in recent studies for a number of countries with low average incomes – Bangladesh, the Arab Republic of Egypt, India, Indonesia, Kenya, Morocco, and Tanzania. The lower limit of the range coincides with the poverty line commonly used in India,*" (World Bank, 1990, p.27). The background work for this analysis is a working paper by Ravallion, Datt, van de Walle and Chan (RDVC) (1991) – an abbreviated version of which appears as Ravallion, Datt, and van de Walle (1991) without the important information on the underlying poverty lines. RDVC (1991, Appendix 1) lists 31 pov-erty lines, from both rich and poor countries, all expressed in dollars per person per month in 1985 international currency. The sources are sometimes World Bank reports, and while some were no doubt created within the Bank, or with Bank assistance, many (perhaps most) of the lines have a genuine local provenance. The lowest of the lines is $23.00 per person per month for India, followed by $31.00 per person per month for Bangladesh, Indonesia, Kenya, Morocco, Nepal, and Tanzania. The Philippines ($32.25) and Pakistan ($34.45) are a little higher. The cluster at $31 (or $372 per annum or $1.02 per day) is the source of the higher of the two lines in the 1990 WDR, and it was this number that was carried through into subsequent work and discussion. In Chen, Datt, and Ravallion (1994), a monthly line of $30.42 is a focal point: this initially mysterious number is, of course, the monthly equivalent of (exactly) $1 a day. The rhetorical force of this originally serendipitous number has been an important part of its adoption into the mainstream of development discourse.

The next round of the ICP was benchmarked in 1993, and the results made their way into versions 6 of the Penn World Table. When the World Bank came to update its poverty estimates, the Penn results were not yet available, and Chen and Ravallion (2001) use instead the

World Bank's own estimates of PPPs using the ICP data. The Bank uses different index number aggregation formulas than does the PWT, so the numbers are conceptually different even though the underlying price data from the ICP are the same. The Bank also took the opportunity of switching from PPPs for GDP as a whole to PPPs for consumption, a conceptual improvement given that the poverty counts are themselves based on levels of household consumption. This was the first occurrence of an issue that was to arise again after the 2005 ICP, and will arise again in the future, of how to update the global poverty line. Because each round of the ICP involves substantial methodological change, and because there are no ICP price data to make a fully satisfactory link between benchmarks, the new PPPs are simply different from the old PPPs, rather than an update. When a domestic consumer price index (CPI) is rebased, we effectively always have a linking factor that permits us to scale up the new series, or scale down the old one, converting, say, 1985 US dollars into 1993 US dollars. But PPPs are *multilateral* indexes so that the linking across bases will give different answers depending on which country is held constant. In particular, the "obvious" recourse of linking through the US dollar, converting 1985 international dollars to 1993 international dollars using the change in the US CPI from 1985 to 1993, while it gives one answer, is not necessarily the answer that we want. I shall return to this point, and hope to clarify the issue as I go.

Chen and Ravallion (2001) resolve the issue by going back to the poverty lines of poor countries, converting to PPPs using the 1993 PPPs, and selecting a global line from the results. I have not been able to find the actual poverty lines that went into this calculation, nor their value in 1993 international dollars, but Chen and Ravallion say that they are the same lines that were used for the 1990 WDR, as described above. They run a regression of the poverty lines on a quadratic of average *per capita* consumption, and use it to estimate a minimal line which turns out to be essentially identical to a procedure that takes the median poverty line from Bangladesh, China, India, Indonesia, Nepal, Pakistan, Tanzania, Thailand, Tunisia, and Zambia. Of these ten countries, six were included in the original $1-a-day calculations, four (China, Thailand, Tunisia, and Zambia) are new, while four from the original list (Egypt, Kenya, Morocco, and Philippines) are dropped. The new line is $1.08 in 1993 international dollars, compared with $1.34, which is the value that would come from taking the original $1, and scaling up by the US CPI in 1993 relative to 1985, which was 144.5 compared with 107.6. Chen and Ravallion's procedure preserves the spirit of the original calculation, going back to Ahluwalia, Carter, and Chenery (1979) though, as they note, it is also possible to argue for updating using the US CPI. In particular, the audience for the international poverty counts is largely based in the rich world, whose citizens are familiar with the dollar, the value of which is well understood. So when a more accurate ICP revises upward the price levels in poor countries, as happened in the 1993 round (and again in 2005), it is true that poor people in, say, India are living on a smaller fraction of dollar than had been previously erroneously calculated, and since the dollar is the yardstick that people understand, the global poverty count – as perceived by the well-off the world – should go up. Going back to the poverty lines of poor countries, as Chen and Ravallion do, eliminates this effect, and takes us closer to the counts of the poor countries themselves, which are, of course, unaffected by changes in PPPs; the

only thing that changes the counts are changes in the relative PPPs between poor countries themselves. The fact that the global line in 1993 dollars ($1.08) was so close to the global line in 1985 dollars ($1.02), although coincidental, may have caused some to think that little had changed, both figures being "close enough" to $1-a-day. Although it is not my main concern here, I should also note that the changes in relative PPPs between the countries in the poverty count also caused major revisions in the structure of global poverty in the 1993 based over the 1985 based numbers (see Deaton, 2001).

The global poverty line is designed to be an *absolute* line set at the minimal acceptable level for anyone on the planet. If it is to be used to document changes in poverty over time, for example in fulfillment of the Millennium Development Goals, then there is certainly a virtue in keeping the poverty line fixed in real terms, so that we know that poverty is diminishing, not that the standard of poverty is being changed. The trouble is that, in a world of multilateral price indexes, there is no unique or obvious way of doing so. Even if we prefer going back to the country poverty lines over scaling up the US dollar for inflation, one might argue that we should stick to the same countries, or better still, the same poverty lines. Countries tend to increase their poverty lines as they get richer, but the global poverty lines do not have to follow, especially if we think that countries are moving from absolute to relative poverty as they get richer. As it is, the 1993 update, which changed the countries, seems thereby to have changed the standards. Of course, it is much easier to criticize the procedure than to propose a fully satisfactory alternative, a live issue that remains open, and that will have to be faced again after the ICP 2011.

Revisions after the 2005 ICP

The most recent revisions to the global poverty lines were in response to the much-improved ICP 2005 which then presented an opportunity to improve the poverty numbers too. As was the case for the previous revision, Chen and Ravallion (2010) used poor-country poverty lines, converted using the new consumption PPPs, to define a global line. Unlike the previous update, they used a new and revised collection of poverty lines, presented in Ravallion, Chen, and Sangraula (2009). Following procedures similar to earlier ones, they ran international regressions of the poverty lines on *per capita* expenditure levels, and showed that, while poverty lines rise with living standards across countries, the relationship is essentially flat among the very poorest countries, suggesting an irreducible minimum *per capita* consumption level that is a good candidate for use as a global absolute poverty line. There are fifteen countries in the list, which appear in the top panel of Table 1, together with their poverty lines, expressed in *per capita* consumption per day in 2005 international dollars. The mean of these lines, $1.25 per person per day, is the Bank's current global poverty line, and there are estimated to be 1.37 billion people in the world living below that level.

As should be clear by now, there are several other ways of calculating the line given the new PPPs. For example, it is notable that there are only two countries, Nepal and Pakistan, that appear in both the 1993 and 2005 versions. While it is certainly true that, as Asia has grown richer relative to Africa, so that we might expect more African countries to appear in the reference group, the revision has as much to do with earlier data availability as with the changing composition of the poorest group

of countries: Tanzania is the only one of the African countries in the top panel of Table 1 that appears in CDV (1991). Again, it is not clear to what extent these lines are locally owned and debated, as opposed to calculated by the Bank, other international organizations, or external NGOs. The mean and median of the 2005 lines of the original countries ($1.16 and $1.05) are also lower than the mean and median of the newly selected poorest countries, in spite of the inclusion of Thailand and Tunisia in the original group. If the original median method is applied to the original ten countries, the global line would have been $1.05, and there would have been less than a billion poor people in the world, about 400 million less than the current counts.

The other possibility is the "rich country audience" procedure that I argued for above. According to this, we hold the old line of $1.08 in 1993 dollars which, given US CPI inflation of 35 percent, is $1.46 in 2005 prices, which would result in 1.76 billion people being classified as poor. So we have a range of 0.97 billion poor to 1.76 billion poor using different, but defensible, methods for updating the global line.

As was the case with the previous revision, the lines that come from averaging poor country poverty lines are lower than the line that comes from updating the previous line using US inflation. This is because the ICP 2005 revised upward the consumption (and GDP) PPPs for most poor countries, *even for the same year*. The second column of Table 1 shows the ratio of the new PPP to the old PPP, for 2005, where the old PPP is the consumption PPP for local currency relative to US dollars based on the 1993 ICP, and updated, and the new PPP is the 2005 consumption PPP from the 2005 ICP. Except for Uganda and Tajikistan, all of these numbers are greater than one so that, relative to the US, all but those two countries have lower estimated consumption under the new PPPs (or equivalently, relative to each of them except the two, the US has higher estimated consumption). This overall relative impoverishment of the poor countries has no effect on the poverty count – because the global line is set from those country lines, and is reduced by exactly the same amount (see Deaton, 2010) – but it alters the relative PPPs between the poor countries. In particular, the revision is larger in the "new" 15 countries than in the "old" ten countries, which means that, had the "new" countries been used prior to the 2005 ICP, the ratio of "new" to "old" global poverty lines would have been even larger than was the case after the 2005 ICP.

Table 1 *Countries and poverty lines for calculating a global poverty line*

Countries	Poverty line 2005 international $	Ratio of 2005 PPP to 1993 PPP for 2005	Hypothetical current PL in 1993 $
15 reference countries 2005			
Chad	0.87	2.33	2.03
Ethiopia	1.35	1.75	2.37
Gambia	1.48	3.03	4.48
Ghana	1.83	1.23	2.26
Guinea-Bissau	1.51	1.79	2.70
Malawi	0.86	2.86	2.45
Mali	1.38	1.85	2.55
Mozambique	0.97	2.38	2.31
Nepal*	0.87	1.96	1.70
Niger	1.10	1.85	2.03
Rwanda	0.99	1.64	1.63
Sierra Leone	1.69	1.49	2.53
Tanzania*	0.63	1.49	0.94
Tajikistan	1.93	0.91	1.76
Uganda	1.27	0.68	0.86
Mean	1.25 (1.37)	1.82	2.17
Median	1.27 (1.41)	1.79	2.26

Countries	Poverty line 2005 international $	Ratio of 2005 PPP to 1993 PPP for 2005	Hypothetical current PL in 1993 $
10 reference countries 1993			
Bangladesh	1.03	1.41	1.46
China	0.85	2.32	1.98
India	0.90	1.41	1.27
Indonesia	1.07	2.13	2.28
Nepal*	0.87	1.96	1.70
Pakistan	1.67	1.47	2.45
Tanzania*	0.63	1.49	0.94
Thailand	1.89	1.15	2.18
Tunisia	1.35	1.81	2.46
Zambia	1.30	1.20	1.57
Mean	1.16 (1.20)	1.64	1.83
Median	1.05 (0.97)	1.48	1.84 (0.93)

Source: author's calculations.

Notes: *Column 1 is per capita consumption per day in 2005 international dollars, taken from RCS (2010). Column 2 is ratio of 2005 PPP for consumption in 2005 international dollars to 1993 dollars. Column 3 is column 1 multiplied by column 2. The numbers in brackets after the means and medians are the numbers of global poor, in billions. Starred countries appear in both lists.*

The final column shows these old and new poverty lines converted at PPPs for 2005, but now using the PPPs for 2005 based on the updated results of ICP 1993. This helps us to isolate the effect of changing the reference countries from the effect of moving to the new ICP. Comparing the first and third columns, we can see that the ratio of the two lines, which using ICP 2005 was 1.08 for the mean and 1.21 for the median, would have been 1.19 for the mean and 1.23 for the median using 1993 based PPPs. The gap between the counts based on the "new" countries poverty line and the "old" countries poverty line would have been much larger had the revision to the ICP not taken place, or not been incorporated into the poverty estimates (note that the average of the 1993 based poverty lines for the "old" countries is much higher than the old $1.08 line. This is in part because of US inflation between 1993 and 2005 – about 35 percent – but also because the poverty lines have themselves been updated since the $1.08 was set).

The new global poverty count is higher than the old poverty count, in part because of the change in the structure of PPPs, but also because the group of countries was changed, dropping those with low poverty lines, and including new ones with high poverty lines. It should also be noted that the 15 new countries have an average population of only 19.9 million in 2005, compared with an average population of 307.7 million in the 10 old countries. The global poverty counts are dominated

by India and China, where about half of the world's poor live, yet the global poverty line, changes in which throw millions of Indian and Chinese in and out of statistical poverty, is dominated by small African countries, some of which are small indeed: Sierra Leone had less than 6 million, and Guinea-Bissau – whose poverty line gets the same weight as the other 14 countries, and infinitely more than India and China, had less than 1.5 million (for further discussion, see again Deaton, 2010).

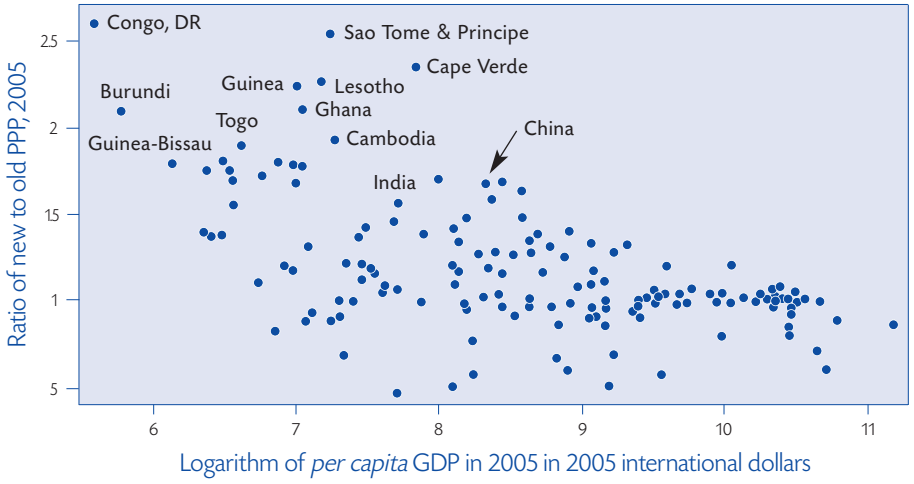
What do we conclude from all this? First, it is not obvious how to maintain a constant poverty line through a new round of the ICP. The Bank's procedures do not do so, causing large shifts in both the structure and total of world poverty. Perhaps the level of global poverty is less important than its rate of reduction. But the level of poverty affects its distribution over countries, and because there are fewer people in Africa than in Asia who live close to the global line, the higher line means a greater "Asianization" of global poverty. Past experience, for example with statistical adjustments that affected urban versus rural poverty, indicates that such adjustments can matter, at least in the debate about who deserves the greater priority. The raising of the line also means that India will no longer meet the first MDG, though the lack of reaction to this change suggests that meeting the MDGs is of largely rhetorical significance, and that accurate measurement is neither here nor there.

1.2.3. The ICP and global inequality

The revisions of the PPP exchange rates in the 2005 ICP generally raised the estimates of price levels in poor countries relative to those in rich countries (see Figure 1). This plots the ratio of new to old PPPs (for GDP) against the new level of the logarithm of *per capita* GDP and shows a significant negative relationship.

Figure 1

Ratios of new (2005-based) PPPs to old (1993-based) PPPs for GDP in 2005

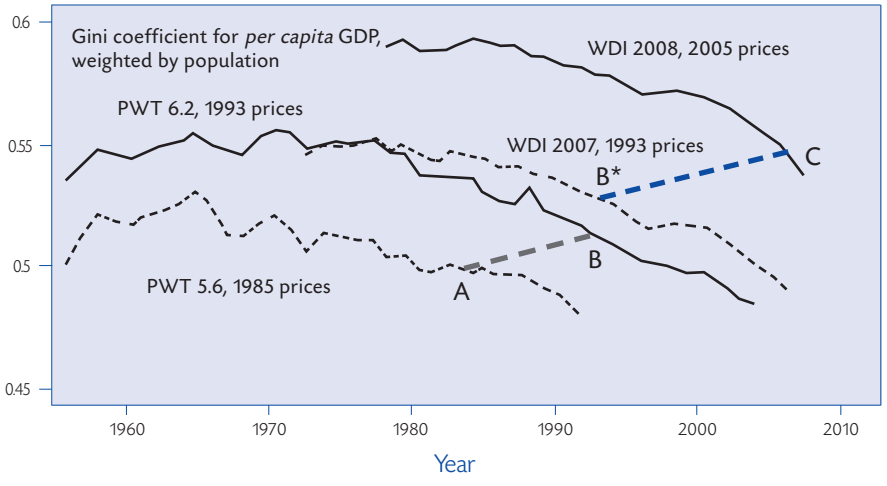


Source: author's calculations, taken from Deaton (2010).

The revision therefore widened the distribution of *per capita* GDP around the world. Figure 2, reproduced from Deaton (2010), shows Gini coefficients for the between country component of *per capita* GDP, and plots both variation over time, and the effects of the last two revisions in the ICP. These measures of income inequality, sometimes referred to as Type II inequality (Milanovic, 2005), take countries as units, and weight each country by its population. They therefore represent the global distribution of income between persons if each person in the world had his or her average country income, and so ignores the contribution to global inequality of within-country inequality. Cross-country inequality is the largest component of total inequality, but within-country inequality has been rising in many (but not all) countries, so that the downward trends in Figure 2 may not be seen in the Gini for the distribution of income between all the persons in the world, the Type III distribution. My concern here is entirely with the cross-country measure and with the effect of successive ICPs on the estimate.

Figure 2

Between country inequality over time, and the effects of successive ICP revisions



Source: Deaton (2010), updated by the author.

Figure 2 shows, using data from the World Development Indicators (WDI), that the substitution of ICP 2005 for ICP 1993 between the WDI 2007 and WDI 2008 resulted in a sharp increase in measured inequality. The World Bank data do not include data from ICP 1985, but the effect of the introduction of ICP 1993 can be seen by comparing inequality using versions 5 and 6 of the Penn World Table. This is shown in the bottom part of Figure 2, and the earlier revision also resulted in an increase in measured inequality. Note that the PWT Gini is lower than the World Bank Gini using the same, ICP 1993, price data.

This is because the PWT uses a Geary-Khamis aggregation procedure, in which the country price indexes are Paasche indexes relative to a global price index, while the World Bank uses a version of the Gini-Eltetö-Köves-Szulc (GEKS) aggregation, in which the underlying indexes are adjusted superlative indexes. The former results in Gerschenkron bias, overstating the incomes of poorer countries, which is avoided by the latter. The essential point here is that, if we compare like with like, PWT with PWT, and WDI with WDI, both of the last two ICP revisions increased measured inequality.

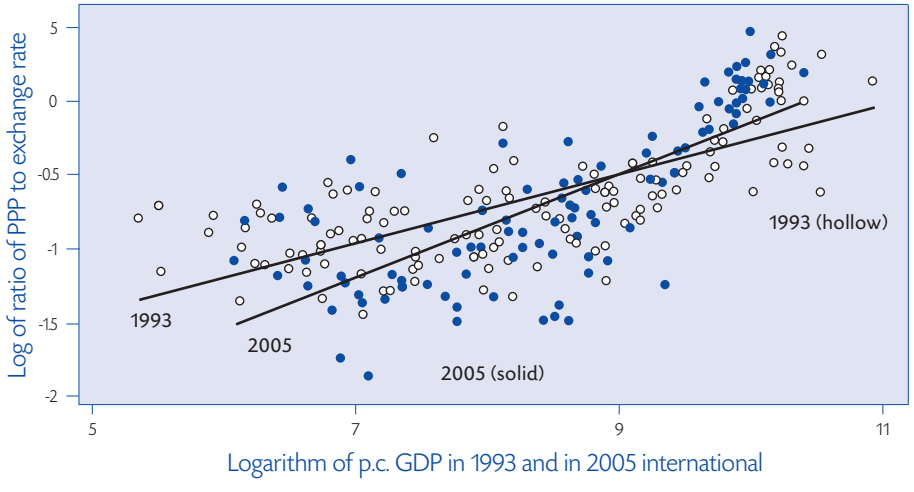
Why do ICP revisions increase the spread of national incomes? One answer is that they do not, at least in general. Each ICP revision has contained a large number of methodological improvements over previous rounds, and these revisions will certainly change measured inequality. In Figure 2, we are looking at only two changes, so that if the effect of the revisions on inequality were unpredictable *ex ante*, there is a one in four chance that both revisions would revise inequality upwards. There have also been a number of substantive explanations put forward, at least about the most recent revision. (I know of no similar work on the revision from ICP 1985 to ICP 1993.) In Deaton (2010), I investigate a number of these; although I do not identify any single factor that can explain all of what happened, there are a number of issues that contributed, including the high prices of some Western goods in poor countries, particularly in Africa, and the fact that goods – such as air travel – which are expensive and rarely used in Africa, attract part of the rich country weight when African countries are compared with rich countries in the multilateral comparisons.

Another possibility, which I did not consider earlier, is that there was no jump in inequality between rounds, and the problem lies, not in the ICP rounds themselves, but in the way that the PPPs are updated between rounds. In terms of Figure 2, using the WDI estimates, the top curve would be correct for 2005, and the middle curve for 1993, but neither curve is necessarily correct for the years in between. If so, the correct assessment of trend would come from connecting the 1993-based estimate for 1993 with the 2005-based estimate for 2005. This gives the inequality trends shown in the figure as AB and B*C, which show international inequality increasing, not decreasing, essentially because the rapid rate of growth of India and China is reduced by introducing the 2005 ICP revisions gradually, instead of all at once. Why might the extrapolation between rounds be problematic? The World Bank updates its PPPs by taking the benchmark PPPs and multiplying by the relative price inflation factors for the country and the US. So that if the benchmark PPP for 2005 is 15 local currency units per dollar, and the local inflation rate to 2010 was 20 percent and that in the US 10 percent (fictionally), the 2010 PPP would be 15 multiplied by the ratio of 1.20 to 1.10, or 16.4. This procedure is an obvious first cut, but is unsatisfactory in a number of ways. One problem is that the content of the domestic CPIs is not matched to the internationally comparable goods and services that appear in the ICP. But perhaps more fundamental is that the procedure ignores one of the main reasons for using PPPs in the first place, which is that the price levels in poor countries tend to be lower in rich countries – the Balassa-Samuelson effect – so that as poor countries grow, we would expect their price levels to rise and this is not taken into account by CPI adjustment, at least not explicitly.

In a recent paper, Ravallion (2010) has made the important argument that the changes in the PPPs from ICP 1993 to ICP 2005 are *not* primarily generated by methodological revisions and improvements, but can be explained by the Balassa-Samuelson effect operating over time, so that countries that have grown more rapidly have seen the largest upward revisions in their price levels. This is an important possibility that, as far as I am aware, has not been previously discussed in the context of ICP revisions. Figure 3 plots the price levels against *per capita* GDP in the two rounds; the data are constructed from the 2007 and 2008 World Development Indicators, but are essentially identical to those used by Ravallion. It shows that the relationship has a steeper slope in 2005 than in 1993 and, as in Figure 1, that the price levels have been revised upwards more in the poorer countries, in addition to any effect that comes from movements along the line. In fact, if we use only the countries that were in both rounds, there is no correlation – or rather an insignificant negative correlation – between the changes in the price levels between the two rounds and the growth of real *per capita* GDP between 1993 and 2005. Ravallion also gets this result, and we differ only in the interpretation that we place on it.

Figure 3

Relationship between price level and GDP in 1993 and 2005



Source: author's calculations.

Consider the equation linking the change in the log price level to the change in log GDP, and write this

(1) $\Delta \ln \pi_i = \alpha + \beta \Delta \ln y_i + u_i$,

where π is the price level, and y is *per capita* GDP in constant international dollars, both for country *i*. This regression has an insignificant estimate of β of -0.11. Note however that the

price level is the ratio of the purchasing power parity exchange rate P to the market exchange rate π. In consequence (1) can also be written in the form

(2) $\Delta \ln \pi_i = \alpha + \beta(\Delta \ln Y_i + \Delta \ln \pi_i) + u_i$,

where Y_i is *per capita* GDP at market exchange rates. (2) then implies that

(3) $\Delta \ln \pi_i = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} \Delta \ln Y_i + \frac{u_i}{1-\beta}$,

The regression (3), using again only the countries included in both 1993 and 2005, yields a positive and significant estimate of $\beta/(1-\beta)$ and thence of β.
Ravallion accepts this estimate as evidence of a Balassa-Samuelson effect operating over time, and rejects the lack of correlation in (1) on the grounds that (1) is contaminated by a negative simultaneity feedback from the growth of the price level (or the real exchange rate) to the growth rate of *per capita* GDP. This is possible, but there are other possible interpretations.

GDP in international dollars is calculated by dividing GDP in local currency by the PPP or, equivalently, by dividing GDP at market exchange rates by the price level estimated from the ICP. Thus Y and π are independently measured, given the reasonable assumption that the exchange rate is accurately measured. In consequence, measurement error in the PPP – which is certainly present – will bias downward the estimate of β in (1), but not the estimate of $\beta/(1-\beta)$ in (3) which, like Ravallion's explanation, could allow (3) to be interpreted in favor of the Balassa-Samuelson effect. However, we might also argue that there is no such effect (or at least that it is too small to detect), that $\beta = 0$, that $\Delta \ln \pi_i$, the changes in the price levels, are driven by methodological and statistical improvements that are unrelated to the growth of GDP, or any other real economic variable, and that the significance of (3) comes from the fact that $\Delta \ln \pi_i$ appears on both the left and right-hand sides of the equation. In consequence, the significance of (3) is not strong evidence for Balassa-Samuelson effects over time. Even so, the significance of (3) *does* show that the growth of *per capita* GDP at market exchange rates has predictive power for the change in the price level over successive rounds of the ICP, either because the growth of GDP at market exchange rates has the growth in the price level as one of its components, or because both are related to other factors, the most obvious being changes in the prices of particular commodities, such as oil, or staples. It also suggests that the revisions are not entirely due to methodological changes between rounds. Since these effects are not taken into account in the updating of the PPPs between rounds, then at least some of the increase in inequality can reasonably be attributed to the failure to do so. In other words, between-country inequality has not been falling as rapidly as we thought. The same would be true of global poverty, were it measured relative to a fixed international dollar. As it is, there is no such effect, or at least it is small because it depends only on revisions to relative PPPs between poor countries.

1.3. Measuring hunger

1.3.1. FAO hunger estimates

The first of the Millennium Development Goals, the elimination of poverty and hunger, has three targets. The first is to halve, between 1990 and 2015, the number of people living under $1 a day. The second is about providing full employment and decent work to all. The third is to halve the number of people living in hunger. But how we measure hunger is as difficult and contentious as the measurement of poverty. The numbers that are usually quoted are provided by the Food and Agriculture Organization (FAO) of the United Nations, and are published annually in their annual report on *The State of Food Insecurity in the World*, the most recent of which is for 2009. However, a September 2010 press release gave the headline number for 2010, that there are 925 million people undernourished which is a decline from 1,023 million in 2009. These numbers measure *undernourishment*, the number of people whose food intake is less than their needs, rather than malnutrition, which measures anthropometric or medical outcomes, including those that are the consequence of undernourishment, for example by being too thin or too short. The FAO calculates undernourishment by calculating total food supplies for each country, converting them to calories, and distributing them over people assuming a log normal distribution, whose variance is estimated from household survey data on calorie consumption. Current estimates, including those for 2009 and 2010, are based on projections of food supplies, since there are no available surveys or food supply data for those years; indeed, at the time of writing, 2010 has some months to run. That the FAO should be able to provide such up-to-date numbers has fueled critical discussion, in particular on the Aid Watch Blog (Easterly, 2010), which also contains a response by David Dawe of the FAO, and by Richard King of Oxfam who provides an excellent summary of the FAO methodology.

One persistent concern about the hunger estimates, like the poverty estimates, is that they are not subject to the checks and balances that surround important national statistics, such as unemployment rates or consumer prices indexes, whose production is insulated from the agencies responsible for policy-making, *e.g.* the central bank or the finance ministry. Publication of the hunger numbers is often accompanied by calls for more aid, although not usually by evidence that more aid would be effective in reducing hunger. I do not believe that the hunger (or poverty) estimates are constructed in anything other than a thoroughly professional way, but I do think that these numbers would be more credible were they subject to better international control, for example by a panel of international statisticians, demographers, or economists.

Beyond the political economy, there are many reasons to question the FAO hunger estimates. In particular, calorie intake is not the same thing as the lack of physical and cognitive functioning that can be threatened by inadequate diet, but which is determined by other factors too, particularly by the disease environment and by the calorie demands of work. It is *net* nutrition, the nutrition that is retained by the body after

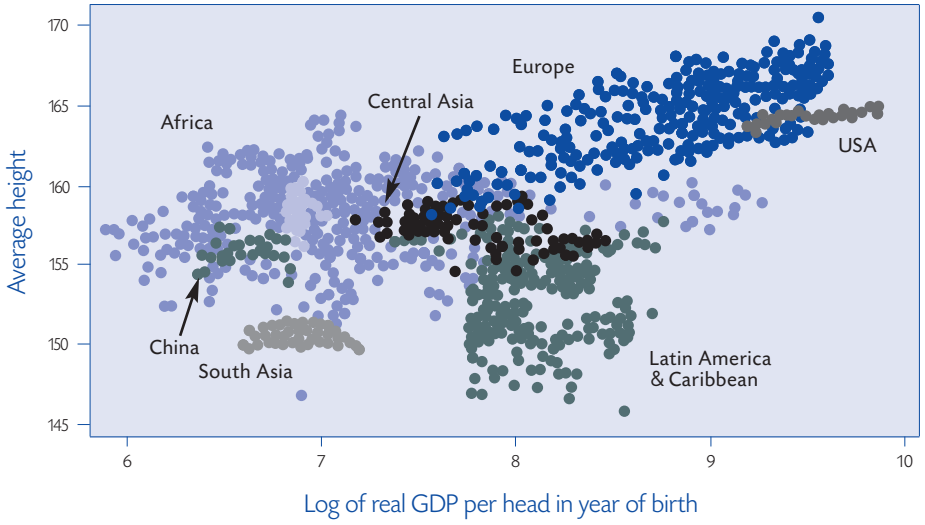
meeting the demands of disease and work, not *gross* nutrition – the intake of food – that affects physical and mental health and the growth and development of children. Of course, it is not a good thing to be hungry, or to get fewer nutrients than are needed, though the measurement of need by fixed cutoffs will often be too crude to be useful. These points are forcefully made by Peter Svedberg (1999) who also notes that calorie-based measures come from *household* consumption surveys, and so cannot yield measures of deprivation for *individuals*. He lists a number of other problems with the FAO procedure, including the inaccuracy of the underlying data, especially in sub-Saharan Africa, the sensitivity of the counts to small variations or small errors in the underlying assumptions, including the calorie cutoffs, and the fact that the hunger counts are almost perfectly predicted by aggregate food availability, leaving little role for local variations in needs or in the distribution of calories over people. In effect, this close link between hunger and total food availability means that the international variation in the hunger measures is dominated by international variation in *per capita* GDP.

1.3.2. Measuring malnutrition: Africa versus Asia

Direct measures of malnutrition do not always follow national income. Svedberg notes that the 1992 FAO counts list Africa as much hungrier than Asia, and this remains true in the most recent counts, for 2004–6, which list 30 percent of people in sub-Saharan Africa as undernourished, as opposed to only 23 percent in South Asia, and 22 percent in India. Yet Africans are generally better nourished than Asians. Figure 4, which is an extended and updated version of Figure 4 of Deaton (2007), plots the average adult heights of birth cohorts of women against *per capita* GDP in the years of their birth; African women are generally taller than Indian, Bangladeshi, and Nepali women (marked as South Asia), in spite of the much lower incomes (and higher FAO hunger estimates) in many African countries. Adult heights are a good indicator of early childhood (net) nutrition, and although it is true that the well-nourished and richer Europeans and Americans are the tallest in the world, there is no relationship between adult height and *per capita* GDP at the time of birth in the rest of the world. Indeed, it remains unclear what does determine the patterns shown in Figure 4. Africans are not only tall on average, but they show enormous dispersion in height from place to place, perhaps because the patterns of nutrition and of disease vary a great deal from country to country, and sometimes even within countries. Beyond that, although Africans typically show less malnutrition than South Asians, they have much higher rates of infant and child mortality, a contrast that is sometimes referred to as the Asian/African paradox (Klasen, 2008). This is a genuine puzzle that is not well understood (although it is possible that the disease environment is worse in Africa, and the nutritional environment better). That the FAO hunger numbers do not solve the paradox is not surprising, and lower malnutrition in Africa does not necessarily imply that lower hunger figures there are wrong, only that malnutrition and hunger are two different things.

Figure 4

Heights and per capita GDP of cohorts of women



Source: author's calculations updated from Deaton (2007).

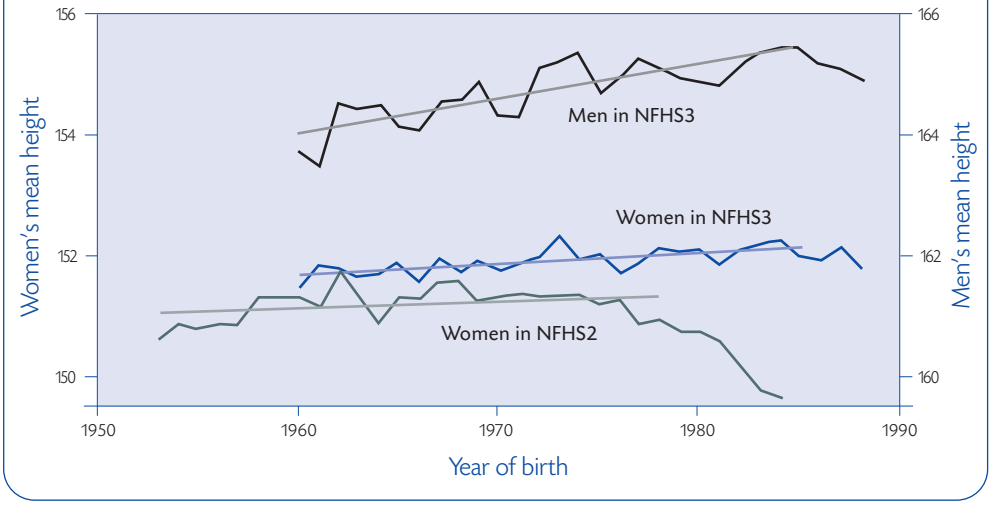
1.3.3. Calories and nutrition in India

That calorie intake and nutrition are not the same is well-illustrated by the situation in India, recently studied by Deaton and Jean Drèze (2009) (DD). Recent economic growth in India has been high by any standards, and markedly so relative to Indian history. Although the reduction in measured poverty is a good deal less than would be warranted by such growth (largely because of the inconsistency between the surveys and the national accounts, though there has also been some increase in inequality), even the poorest groups have seen real progress. Yet *per capita* calorie consumption has been *falling*, especially in rural India where *per capita* calorie consumption fell by about ten percent from 1983 to 2004–5. The reduction in calories from cereals – the basic Indian staple – has fallen more rapidly than the total, by about 300 calories per person per day in rural India, and about half as much in urban India. While there has been a long-term decline (60 years) in the consumption of "coarse" grains – sorghum, millet, and maize – *per capita* rice consumption has been falling for 20 years, and *per capita* wheat consumption has been more or less constant for the last decade. Given these numbers, if we use an FAO method to calculate the number of those in hunger, here defined by people who live in households whose *per capita* calorie consumption is less than 2,100 calories per day, we find that 76 percent of the Indian population was hungry in 2004–5, compared with "only" 65 percent in 1983 (DD, Table 5). (Note that we are currently awaiting a new large consumption survey for India, so that DD do not include years beyond 2005 in their analysis.)

In contrast to the calorie decline, direct measures of malnutrition, for both adults and children, show improvement over the same period, albeit at a rate that is slower than desirable, and without challenging India's place as one of the most malnourished countries on earth. DD review the (often incomplete) estimates of malnutrition. From the mid-1970s to 2005, these estimates show declines in the fractions of children who have low weight for height, low weight for age, and low height for age, as well as reductions in clinical signs of malnutrition. Yet 46.7 percent of Indian children are still too light for their age, percentages that are exceeded only by children in Nepal and Bangladesh, with Timor-Leste, Yemen, Burundi, Madagascar, Sudan, Laos, Niger, Eritrea, and Afghanistan completing the list of the ten worst countries (DD, Table 10). Adults are also doing better, at least if we again judge by their heights as adults. Figure 5, reproduced from DD, shows the heights of adult men and women by birth cohort, taken from the two most recent National Family Health Surveys (the Indian DHS). NFHS2 collected data in 1998–99, but only on the heights of women, while NFHS3, which collected data in 2005–06, measured both men and women. The Figure shows that later born women and men (shown on a different scale) are taller (except for those on the right, who are not fully grown), indicating a clear improvement in nutrition over time. (Note that there are some inconsistencies of measurement in women's heights between the two surveys.) Yet once again the situation is far from uniformly positive. Men are getting taller at about three times the rate at which women are becoming taller. While we do not know why this is the case, it is unlikely to be differences in calorie intake – for which there has never been any evidence – and in any case, the calorie-based measures cannot distinguish between men and women because they use household-level data. Even among men, the rate of improvement is about half the rate of improvement in China, where there is no difference in progress between men and women. The rate of progress in China is about the same as it was in Europe and the United States since World War II. Interestingly, this rate of growth is reproduced in Kerala in India, where there is also no difference between men and women. (Tamil Nadu is not far behind.) Yet *per capita* calorie consumption in Kerala and Tamil Nadu is amongst the lowest in India.

Figure 5

Heights of Indian men and women, by birth cohort



Source: Deaton (2010), updated by the author.

We do not know exactly why calorie consumption and malnutrition are so different across space and time in India. The leading hypothesis is that there has been a reduction in heavy manual labor, which has reduced the need for calories for fuel. Greater mechanization of farm labor is one reason; others include a huge improvement in roads – so that people do less walking and less carrying of heavy loads–and better provision of water – reducing the need for carrying water over long distances. The improvement in water provision may also have reduced the prevalence of water-borne disease, and the calorific toll that it exacts. While there is little or no *direct* evidence for these explanations, they are consistent with much of the evidence – for example that the higher wage states are those with lower *per capita* calorie consumption, and the same temporal reduction in calories appears to be occurring in China too. In any case, if reductions in calorie intakes reflect reductions in need – even at a time when malnutrition is stunningly high – we cannot use calorie based measures to estimate the prevalence of hunger, either over space or over time.

The obvious alternative is to use the anthropometric measures directly. Here there has been enormous progress, through the spread of the Demographic and Health Surveys. These have greatly extended their measurement of height and weight, first to children, then to women of childbearing age, and most recently – though there are still only a few surveys – to men. These surveys are as close to a gold standard as we are going to get in this area, although the irregularity of the DHS surveys makes it difficult to use them for monitoring, for example for assessing the effects of the food price crisis on the heights and weights of children.

1.3.4. Asking about not having money for food: the Gallup World Poll

There is one other possible way of measuring hunger, to which I devote the remainder of this section. In the Indian National Sample Surveys, respondents are asked questions about hunger. The Indian questions are "*Do all members of your household get two square meals a day?*" with answers of "yes", "in some months", or "no". In the most recent surveys, the "two square meals a day" has been replaced by "enough food every day". The answers to those questions, unlike the calorie questions, but like the malnutrition numbers, show a steady improvement over time, albeit with a good deal of variation across states; over all India, the fraction of households responding other than "yes" declined from 17.3 percent in 1983 to 2.5 percent in 2004-05. These questions are cheap to ask, and respondents appear to have no difficulty in answering them. They are therefore likely to be useful for monitoring, especially in the short-run, and until the anthropometrics from the DHS become available.

The Gallup Organization includes a hunger question in its World Poll, which started in 2006, and which has to date (September 2010) collected data, using an identical questionnaire, in 155 countries. Although not all countries are included every year, most countries appear in multiple years; there were 129 countries in 2006, 100 in 2007, 124 in 2008, 118 in 2009, and at the time of writing there are data from 31 countries in the 2010 round. The question is "*Have there been times in the past 12 months when you did not have enough money to buy food that you and your family needed?*" Most of the countries have sample sizes of about 1,000, so that for a yes/no question, the standard error of the fraction reporting yes is $\sqrt{p(1\text{-}p)/1000}$, which if $p = 0.4$, say, would be 0.015, or perhaps twice that if we allow for the design effect.

Table 2 *Fractions of population reporting that they did not have enough money for food (selected countries)*

	2006	2007	2008	2009	2010
East Asia					
China	0.37	-	0.16	0.17	-
Indonesia	0.29	0.25	0.22	0.23	0.25
Philippines	0.60	0.64	0.59	0.68	0.62
South Asia					
India	0.35	0.26	0.23	0.29	-
Pakistan	0.33	0.26	0.28	0.34	-
Bangladesh	0.25	0.24	0.27	0.23	0.29
Sub-Saharan Africa					
Nigeria	0.58	0.56	0.55	0.60	-
Ethiopia	0.27	0.39	-	-	-
South Africa	0.45	0.48	0.56	0.55	-
Kenya	0.73	0.56	0.68	0.63	0.57
Latin America					
Brazil	0.20	0.21	0.21	0.20	-
Mexico	0.36	0.28	0.33	0.34	-

Source: author's calculations.

Notes: *author's calculations from the Gallup World Poll. The 2010 data were incomplete at the time of writing. The question is "*Have there been times in the past twelve months when you did not have enough money to buy food that you or your family needed?*"*

Table 2 shows the fractions of the population reporting this kind of hunger for a number of selected large countries in four of the World Bank's standard regions, East Asia and the Pacific, South Asia, sub-Saharan Africa, and Latin America and the Caribbean. The dots show years where there are no data, mostly for the incomplete 2010 survey, but also where the country was not included, here China in 2007, and Ethiopia in 2008 and 2009. In most cases, the year-to-year variation is small enough to be within the bounds of credibility, though there are exceptions, including China in the first year; I drop this observation in the imputations that follow. For the selected countries, Africa shows more hunger than Asia, which suggests that these measures, like the FAO numbers, are closer to income numbers than the malnutrition numbers.

Table 3 *Estimated numbers of people with not enough money for food (millions)*

	2006	2007	2008	2009
All countries				
Low income	972	808	804	890
Low middle	551	576	592	599
High middle	185	150	176	177
High income	117	89	90	113
World	*1825*	*1623*	*1662*	*1779*
Low and low middle income				
East Asia	383	404	365	400
Europe & Central Asia	50	54	50	55
Latin America	108	110	113	113
Middle-East & N. Africa	78	78	102	67
Sub-Saharan Africa	409	364	410	411
South Asia	511	400	367	439
Total	1539	1410	1407	1485

Source: author's calculations.

Notes: calculated by regressing the fractions of people reporting not enough money for food on country and year dummies separately by income group. The predictions of the regressions are used to fill in missing values and totals are calculated by multiplying the predicted fraction for each country by population and summing over the income group. Because the imputations are done differently in the bottom than in the top panel, the sum of low income and low middle income in the top panel is not the same as the total in the bottom panel. Author's calculations from Gallup World Poll. See Notes to Table 2 for the underlying question.

Table 3 attempts to turn the country estimates into world counts of the total number of people reporting that they did not have enough money for food. Given that some countries are missing in some years, it makes no sense to add up the total numbers in the surveys, because the year-to-year variation will then be affected as much by the selection of countries – in 2007, China is absent–and there would be a large drop in the number of people reporting hunger. Instead, I have filled in the missing values from a simple factor model in which I first aggregate up to the country/year level, and then regress the log of the fraction reporting hunger on a set of year and country fixed effects; the results are the same if I use the fractions themselves instead of their logs. In the top panel of Table 3, which looks across the World Bank's income classifications, the factor regression is done separately for all countries within an income class, so I am assuming that the year-to-year variations around the country fixed effect is the same for all countries within each income grouping. In the second panel, I drop all of the high middle income and high income coun-

tries, and then split up the six remaining World Bank regions, and re-estimate the factor model for each. As a result, in the bottom panel, the time variation is the same within regions, but not across them. In both panels, when I have real data on the fraction hungry I use it, and when not, I use the appropriate factor for imputation. In all cases, the fractions are converted to totals by multiplying by the population.

The absolute size of these numbers is of little importance, and will certainly vary with the precise wording of the question. More important is that the Gallup data confirm a substantial increase in the number of hungry people from 2008 to 2009, by 117 million worldwide, and 78 million in low and low middle income countries. As argued by the FAO, such an increase is entirely plausible given the food price spike in 2008 and the financial crisis that began in that year. Of the worldwide increase, most is in low income countries, though there was also a substantial increase, form 90 to 113 million in high income countries. In the low and middle income countries, the increase is entirely attributable to increases in South and East Asia, with no increase in Latin America, Europe and Central Asia, or sub-Saharan Africa. Indeed, by 2009 there are more hungry people in South Asia than in sub-Saharan Africa, though the fractions are twice as high in Africa. The data actually show a decrease in those reporting hunger in North Africa and the Middle-East; a good deal of this is imputation (Iran, Morocco, and Yemen), but all of the large countries for which there are actual data (Algeria, Egypt, Iraq, Tunisia) show a reduction between 2008 and 2009.

The Gallup estimates show no evidence of an increase in hunger from 2005/7 to 2008, unlike the FAO who show almost as large an increase over this period as between 2008 and 2009. Instead, the Gallup data show what looks like a steady improvement until the year after the financial crisis and the food price increase.

1.4. From many to one: single indexes of development

Although there are good theoretical arguments against attempting to combine indicators in different dimensions (see for example Sen, 1999; Broome, 2001), there is always pressure to construct a single index that can be used to rank countries and to measure progress over time. The UNDP's Human Development Index (HDI), which combines (country aggregate) measures of health, literacy, and income, is perhaps the best known of these indicators. Such indexes present no theory to justify the method of combination (or the weighting of the components), so they have a large component of arbitrariness. Even so, they have the advantage that they recognize the correlation between different dimensions of well-being and deprivation. Countries with low GDP *per capita* also tend to have low life expectancy and low literacy, so that an index number that combines them will give a better picture of the gulf between poor countries and rich countries than does income alone. However, because the HDI uses only national averages, it ignores the correlation between deprivations within countries, that poor Indians are more likely to be sick and less educated. The new multidimensional indexes (Alkire and Santos, 2010) are an ambitious attempt to address this gap. Their measure combines poverty in several dimensions at the *household* level, which solves the within-country correlation problem, at the price of the heavy data requirement that *all* indicators that vary across *households* must be available from the same survey. That such indicators can be computed at all – Alkire and Santos use the DHS surveys, backed up by the MICS and WHS surveys – is an eloquent testimony to the extraordinary enrichment of the data environment in recent years.

Economics has a theory – albeit not a very good theory – of how to combine health and income. It is the same theory that is used to construct measures of the value of life. In the simplest version, consumers are assumed to maximize the lifetime sum of each period's utility, which is itself a function of each period's consumption. Additional years of life add more periods in which consumption can take place (just as additional time at pasture makes fatter cows), so that any given increase in years of life can be turned into its money equivalent, defined as the amount of additional money that would give the same increase in lifetime utility. The simplest version of this sort of accounting is to multiply income by life expectancy, although more sophisticated versions have been proposed (and taken to the data) by Becker, Philipson and Soares (2005), and more recently by Jones and Klenow (2010). These procedures "solve" the arbitrary weighting problem in the HDI, provided, of course, that the theory is acceptable on other grounds. Like the HDI, and because of the correlation between income and health, these measures show much more global inequality in "full" income (which includes a value for life expectancy) than in (regular) income. For most of the post-WWII period, and until the HIV/AIDS

epidemic, life expectancy rose more rapidly in poor countries than in rich countries, so that "full" income inequality declined faster than income inequality, though this was reversed with the dramatic reductions in life expectancy in the AIDS affected countries.

There are a number of problems with these calculations. To assume that African lives are worth less than American or European lives simply because they consume less adds insult to injury. Not only do you get less than I do, but because of that, you yourself are worth less than I am; there is more to life than consumption, and people are not cattle being fed for non-cattle related ends. Beyond that – if more is indeed needed – life expectancy, which may seem innocuous, also contains an implicit aggregation that is problematic. Life expectancy is an aggregate of age-specific mortality rates, but it is one *specific* aggregate among many possible aggregations. In particular, the increase in life expectancy in poor countries has largely been driven by declines in infant and child mortality, whereas the increase in rich countries has come from declines in adult mortality, particularly from declines in mortality from cardiovascular disease and lung cancer. Life expectancy gives very high weight to lives saved at the beginning of life, and relatively little to saving the lives of 50-year-olds. While there is no agreement on which should be weighted more highly, it is far from clear that the life-expectancy weighting is the right one to choose. Reductions in the mortality rates for very young children are, at least to some extent, later accompanied by compensating reductions in fertility by parents. If so, the age structure of the population may not change very much in response to the reduction in mortality, with children who would have died soon after birth "replaced" by fewer children ever being born. There is a clear welfare gain to the parents who do not have to live through the deaths of their young children, and to women who have gone through fewer pregnancies, but those gains are hardly measured by life expectancy. If there is anything to this argument, the narrowing of the life expectancy gap between rich and poor countries from 1950 into the 1980s is not a good measure of decline in inequality.

If the relationship between income and health were sufficiently strong, we might not need to consider both, but make do with one, and let the other look after itself. Perhaps either GDP *per capita* or life expectancy can serve as an index of development? This argument appears in a number of forms. One is what might be called "income fundamentalism", that if countries experience sufficient economic growth, then health will look after itself, perhaps the best statement of which is Pritchett and Summers (1996). Another recent argument comes in a paper by Young (2010), who correctly notes that the data on growth and GDP from Africa are highly unreliable, so that we actually know very little about growth in Africa over recent decades. But there have been substantial improvements in other indicators, including health and mortality of children, from which Young infers that African growth has been much higher than is shown by their national income statistics.

One weakness in both of these arguments is that the correlation between growth and health improvement is very far from perfect, and that the divergence between the two is of considerable interest in its own right, so that we lose a great deal by ignoring it, or by treating it as *entirely* measurement error. The international relationship between life

expectancy and *per capita* income – the Preston curve – is certainly strong, but there are many exceptions where countries have managed to have good health at low income or to have poor health at high income, and at least some of this is explained by policy, not by measurement error. Nor did the policy always have to wait for economic growth. Many lives were saved by vector control in the years following World War II, even in countries with low income and negligible economic growth: some (although certainly not all) health innovations are cheap, and can be put in place even in otherwise unhelpful environments. Yet another example comes from India and China. Up until the Chinese economic reforms, when growth was relatively weak by subsequent standards, infant and child mortality declined rapidly. Afterwards, as resources were switched into production, with health relatively neglected, the progress in infant and child mortality slowed or halted. Meanwhile, progress in India was more gradual, and in spite of its lower overall rate of economic growth, infant and child mortality rates are now close to catching up with China, and have more than caught up in parts of the country (Drèze and Sen, 2002, chapter 4).

Within India, the rate of decline of infant mortality has declined somewhat in the face of more rapid economic growth. Finally, in Deaton (2007), I show that while the cross-country correlation between economic growth and the proportional rate of decline in infant mortality is (as expected) *negative*, there is a small *positive* correlation between economic growth and absolute declines in infant mortality. This happens because the proportional rates of decline in infant mortality have been higher in the richer countries, even from low initial levels, and because rich countries have typically grown faster than poor ones – the well-known divergence in country income levels. The underlying correlation here is not a change-on-change correlation, but a correlation between income *growth* and the *level* of infant mortality. The literature contains a number of possible explanations, but at least one possibility is that good governance contributes to both; if so, we have another case where it is policy that drives at least some of the difference between income and health. If we confound them, we lose out, both on measurement and the understanding of mechanisms.

References

Ahluwalia, M.S., N.G. Carter and H.B. Chenery (1979), "Growth and poverty in developing countries", *Journal of Development Economics*, 6: 299-341.

Alkire, S., and M.E. Santos (2010), "Acute multidimensional poverty: a new index for developing countries", OPHI Working Paper No.38, July.

Becker, G., T.J. Philipson and R.R. Soares (2005), "The quantity and quality of life and the evolution of world inequality", *American Economic Review*, 95(1): 277-91.

Bhalla, S.S. (2002), *Imagine there's no country: poverty, inequality, and growth in an era of globalization*, Institute for International Economics. Washington, D.C.

Bourguignon, F., and C. Morrisson (2002), "Inequality among world citizens: 1820-1922", *American Economic Review*, 92(4): 727-44.

Broome, J. (2002), "Measuring the Burden of Disease by Aggregating Well-being", Chapter 3.1 in C. J. L. Murray, J. A. Salomon, C. D. Mathers and A. D. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*, World Health Organization, Geneva.

Chen, S., G. Datt and M. Ravallion (1994), "Is poverty increasing in the developing world?" *Review of Income and Wealth*, 40(4): 359-76.

Chen, S. and M. Ravallion (2001), "How did the world's poorest fare in the 1990s?" *Review of Income and Wealth*, 47(3): 283-300.

Chen, S. and M. Ravallion (2010), "The developing world is poorer than we thought, but no less successful in the fight against poverty", *Quarterly Journal of Economics*, 125(4): 1577-1625.

Deaton, A. (2001), "Counting the world's poor: problems and possible solutions", *World Bank Research Observer,* 16(2); 125-47.

Deaton, A. (2005), "Measuring poverty in a growing world (or measuring growth in a poor world)", *Review of Economics and Statistics*, 87(1): 1-19.

Deaton, A. (2007), "Global patterns of income and health: facts, interpretations, and policies", WIDER Annual Lecture, Helsinki, UNU-WIDER.
http://www.princeton.edu/rpds/papers/pdfs/deaton_WIDER_annual-lecture-2006.pdf.

Deaton, A. (2007), "Height, health, and development", *Proceedings of the National Academies of Science*, 104(33): 13232-37.

Deaton, A. (2008), "Income, health, and well-being around the world: evidence from the Gallup World Poll", *Journal of Economic Perspectives*, 22(2): 53-72.

Deaton, A. (2010), "Price indexes, inequality, and the measurement of world poverty", *American Economic Review*, 100(1): 5-34.

Deaton, A. and J. Drèze (2009), "Food and nutrition in India: facts and interpretations", *Economic and Political Weekly*, 44(7): 42-65.

Deaton, A. and O. Dupriez (2011), "Purchasing power parity exchange rates for the global poor", *American Economic Journal: Applied*, 3 (1): 137-166.

Deaton, A. and A. Heston (2010), "Understanding PPPs and PPP-based national accounts", *American Economic Journal: Macroeconomics*, 2(4): 1-35.

Drèze, J. and A.K. Sen (2002), *India: development and participation*, Oxford University Press, N. Delhi.

Easterlin, R.A. (1974), "Does economic growth improve the human lot? Some empirical evidence", in R. David and M. Reder (eds), *Nations and households in economic growth: essays in honor of Moses Abramowitz*, 89-125, Academic Press, New York.

Easterly, W. (2010), "Spot the made-up world hunger numbers", Aid-Watch Blog, September 15, 2010, http://aidwatchers.com/2010/09/spot-the-made-up-world-hunger-numbers/.

Filmer, D. and L.H. Pritchett (2001), "Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in states of India", *Demography*, 38(1): 115-32.

Food and Agriculture Organization of the United Nations (2009), *The state of food insecurity in the world: economic crises – impacts and lessons learned*, FAO, Rome.

Jones, C.I. and P.J. Klenow (2010), "Beyond GDP? Welfare across Countries and Time", NBER Working Paper No.16352, National Bureau of Economic Research, Inc.

Kahneman, D. and A. Deaton (2010), "High income improves evaluation of life but not emotional wellbeing", *Proceedings of the National Academies of Science*, Published online September 7th.

Kahneman, D. and J. Riis (2005), "Living and thinking about it: two perspectives life", in F. Huppert, N. Baylis, and B. Keverne (eds), *The Science of Well-being*, 285-304, Oxford University Press, Oxford.

Klasen, S. (2008), "Poverty, undernutrition, and child mortality: some inter-regional puzzles and their implications for research and policy", *Journal of Economic Inequality*, 6(1): 89-115.

Kravis, I.B., A.W. Heston, and R. Summers (1978), "Real GDP *per capita* for more than one hundred countries", *The Economic Journal*, 88: 215-42.

Kulshreshtha, A.C. and A. Kar (2005), "Estimates of food consumption expenditure from household surveys and national accounts", Chapter 7 in A. Deaton and V. Kozel (eds), *The Great Indian Poverty Debate*, 102-18, Macmillan, New Delhi.

Kuznets, S. (1955), "Economic growth and income inequality", *American Economic Review*, 45(1): 1-28.

Layard, R. (2005), *Happiness: Lessons from a New Science*, Penguin, London.

Milanovic, B. (2005), *Worlds Apart: Measuring International and Global Inequality*, Princeton University Press, Princeton.

Pinkowskiy, M. and X. Sala-i-Martin (2009), "Parametric Estimations of the World Distribution of Income", NBER Working Paper No. 15433, October.

Pritchett, L. and L.H. Summers (1996), "Wealthier is healthier", *Journal of Human Resources*, 31(4): 841-68.

Rajaratnam, J.K., J.R. Marcus, A. Levin-Rector, A.N. Chalupka, H. Wang, L. Dwyer, M. Costa, A.D. Lopez and C.J.L. Murray (2010*a*), "Worldwide mortality in men and women aged 15-59 years from 1970 to 2010: a systematic analysis", *The Lancet*, 375: 1704-20.

Rajaratnam, J.K., J.R. Marcus, A.D. Flaxman, H. Wang, A. Levin-Rector, L. Dwyer, M. Costa, A.D. Lopez and C.J.L. Murray (2010*b*), "Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress toward Millennium Development Goal 4", *The Lancet*, online publication, May 24: 1-21.

Ravallion, M. (2010), "Price levels and economic growth: making sense of the PPP revisions between ICP rounds", Policy Research Working Paper 5229, Washington, D.C. The World Bank, March.

Ravallion, M., S. Chen and P. Sangraula (2009), "Dollar a day revisited", *World Bank Economic Review*, 23(2): 163-84.

Ravallion, M., G. Datt, D. Van de Walle and E. Chan (1991), "Quantifying the Magnitude and Severity of Absolute Poverty in the Developing World in the Mid-1980s", Working Papers WPS 587, World Bank, Washington, D.C.

Ravallion, M., G. Datt and D. van de Walle (1991), "Quantifying absolute poverty in the developing world", *Review of Income and Wealth*, 37(4): 345-61.

Sala-i-Martin, X. (2006), "The world distribution of income: falling poverty and c onvergence, period", *Quarterly Journal of Economics*, 121(2): 351-97.

Sen, A.K. (1999), *Development as Freedom*, Knopf, New York.

Sen, A.K. (1987), *On Ethics and Eeconomics*, Blackwell, Oxford.

Stevenson, B. and J. Wolfers (2008), "Economic growth and subjective well-being: reassessing the Easterlin paradox", *Brookings Papers on Economic Activity*, Spring, 1-87.

Stigltiz, J.E., A.K. Sen and J.-P. Fitoussi (2009), *Report of the Commission on the Measurement of Economic Performance and Social Progress*, http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

Summers, R. and A. Heston (1991), "The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988", *Quarterly Journal of Economics*, 106(2): 327-68.

Svedberg, P. (1999), "841 million undernourished?" *World Development*, 27(12): 2081-98.

World Bank (2008), *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program*, The World Bank, Washington, D.C.

Young, A. (2009), "The African Growth Miracle", processed, http://econ.as.nyu.edu/docs/IO/11950/AfricanGrowthMiracle.pdf.

Conférences & Séminaires

2. Multidimensional Poverty and its Discontents

2. Multidimensional Poverty and its Discontents[2]

Sabina Alkire, Oxford University

Abstract

More data on non-income dimensions of poverty are available than at any previous time in history. Alongside this, multidimensional measurement methodologies have advanced considerably in the past fifteen years. These advances together have created new possibilities to measure multidimensional poverty at the local, national and international level. Yet the fact that one can construct an overall measure does not mean that it will necessarily add value. Considering multiple dimensions does not necessarily require a multidimensional poverty index. This paper focuses on the question of when, how and why certain multidimensional poverty measures may add value, sketches the limits of the contribution, and introduces a set of standing questions. The key value-added of a rigorously implemented multidimensional poverty index is that it conveys additional information not captured in single-dimensional measures on the joint distribution of disadvantage and the composition of poverty among different multiply deprived groups. It also provides a consistent account of the overall change in multidimensional poverty across time and space. To makes this case and explore these issues, the paper discusses one general approach to multidimensional poverty measurement – that which reflects joint distribution. It then presents one class of poverty measures within this approach. It also introduces one recent implementation of one measure within this family: the new 100+ country Multidimensional Poverty Index launched in 2010. Pointing to the added value or multidimensional poverty indexes this is not to suggest that single-dimensional measures be abandoned but rather supplemented. Investing further in multidimensional measures has the potential to generate significant advances in understanding and useful policy tools.

[2] I am grateful to François Bourguignon, Stefan Dercon, James Foster, Jeni Klugman, Maria Emma Santos & Gaston Yalonetzky and to the organisers and participants in the 2010 EUDN meetings for comments and suggestions and to Gisela Robles and Ann Barham for research assistance. All errors remain my own.

Introduction

The multidimensionality of poverty is not in dispute.[3] Poverty can mean poor health, inadequate education, low income, precarious housing, difficult or insecure work, political disempowerment, food insecurity, and the scorn of the better-off. The components of poverty change across people, time, and context, but multiple domains are involved.

An emerging question is how multidimensionality should be reflected in measures of poverty.[4] The launch of a new 104-country multidimensional poverty index (MPI) in 2010 attracted attention and interest in many countries, and provoked lively discussion.[5] This paper examines how one aggregate measure of multidimensional poverty adds value to an assemblage of deprivation and income poverty indicators. These issues have become vivid both due to an increasing body of studies on the interrelationships among indicators of disadvantage, as well as to the increased possibility of creating multidimensional poverty measures.[6] More data on non-income dimensions of poverty are available than at any previous time in history. Alongside this, multidimensional measurement methodologies have advanced considerably in the past fifteen years. These advances together have created new possibilities to measure multidimensional poverty at the local, national and international level. Yet the fact that one can construct an overall measure does not mean that it will necessarily add value. As Sen writes, "*The passion for aggregation makes good sense in many contexts, but it can be futile or pointless in others... The [overall] view does have its uses, but it has no monopoly of usefulness. When we hear of variety, we need not invariably reach for our aggregator.*"[7] This paper will focus on the question of when, how and why certain multidimensional poverty measures may add value, sketch the limits of the contribution, and introduce a set of standing questions.

To explore these issues, the paper discusses one general approach to multidimensional poverty measurement – that which reflects joint distribution. It then presents one class of poverty measures within this approach, namely an extension to the FGT class of measures (Foster, Greer, and Thorbecke, 1984) proposed by Alkire and Foster (2007; 2011*a*). It also introduces one recent implementation of one measure within this family: the Multi-

[3] Grusky and Kanbur, 2006; Jenkins and Micklewright, 2007; Ravallion, 1996; Ravallion, 2010; Sen, 1992; Sen, 1993; and Thorbecke, 2008. For example, Cappellari and Jenkins write (in Jenkins and Micklewright, 2007), "*It is widely agreed nowadays that being poor does not simply mean not having enough money*" (opening sentence, p. 166).

[4] In Alkire and Foster (2011*b*), we clarify our measurement methodology and its basis in unidimensional poverty methods; this paper builds upon it, and highlights additional issues in empirical implementation.

[5] After the launch of the Multidimensional Poverty Index (MPI) in July 2010, the Oxfam Blog, the World Bank Blog, VOXEU, and the Journal of Economic Inequality carried substantive exchanges on the MPI.

[6] Of course, there is also a parallel rise in multidimensional welfare assessments. See for example Becker, Philipson and Soares, 2005; Brighouse and Robeyns, 2010; Fleurbaey, 2009; Fleurbaey and Gaulier, 2009; Jones and Klenow, 2010; Kreitler and Kreitler, 2006; McGillivray, 2007; Robeyns and Van der Veen, 2007; Stiglitz, Sen and Fitoussi 2009, among others.

[7] Sen, 1987*b*, p.33.

dimensional Poverty Index. While this paper focuses on one broad approach to poverty measurement, it is important to acknowledge that measurement comprises but a subset of the broad range of techniques that have been developed to assess multidimensional poverty; other methods include qualitative and participatory techniques, dashboards and poverty profiles, dominance techniques, multivariate techniques, and multidimensional inequality indices. Among multidimensional poverty measures, this paper also covers a narrow terrain, and does not address relevant and interesting measures that use information theory,[8] fuzzy set theory,[9] latent variable techniques,[10] multiple correspondence analysis,[11] alternative counting approaches,[12] alternative axiomatic approaches,[13] or dominance.[14] While a number of the research questions are shared among approaches, in the limited space available we can only formulate the issues for one measurement approach. However, it seems possibly useful to set out a clear account of this particular approach, so that its strengths and limitations can be grasped, and areas for further research advanced efficiently.

In the absence of such an account, multidimensional measures of poverty may be viewed as a somewhat sweet distraction. In an eloquent criticism of the parsimony for which economics is known, A. O. Hirschman (1984) proposed complicating economic discourse by, among other things, introducing a more adequate treatment of love. Love, Hirschman argued, is poorly handled in economics, being neither a scarce resource nor an augmentable skill. Lofty as Hirschman's suggestion might have been it did not, in practice, take off. There could be many reasons that parsimony endured in this respect; perhaps it was not sufficiently clear when and how such a complication would add value, or perhaps it has yet to find its time. While multidimensional poverty measurement might seem more familiar to economists than Hirschman's favored topic, it runs the risk of seeming to threaten legitimate parsimony if its potential contribution – and the limits of its contribution – are not sketched more precisely.

[8] Deutsch and Silber, 2005; and Maasoumi and Lugo, 2008.

[9] Balestrino, 1998; Cerioli and Zani, 1990; Cheli and Lemmi, 1995; Chiappero-Martinetti, 1994; Chiappero-Martinetti, 2006; Deutsch and Silber, 2005; Lelli, 2001; Lemmi and Betti, 2006; and Qizilbash 2002.

[10] Kakwani and Silber, 2008*b*; Krishnakumar, 2004; Krishnakumar and Ballon, 2008; and Schokkaert and Van Ootegem, 1990.

[11] Asselin, 2009.

[12] Atkinson, 2003; Erikson, 1993; Gordon, Nandy, Pantazis, Pemberton and Townsend, 2003; Nolan and Whelan, 1996; and Subramanian, 2007.

[13] Bossert, D'Ambrosio and Peragine, 2007; Bourguignon and Chakravarty, 2002; Chakravarty, 1998; Chakravarty and D'Ambrosio, 2006; Chakravarty and Silber, 2008; Deutsch and Silber, 2005; and Tsui, 2002.

[14] Duclos, Sahn and Younger, 2006.

2.1. Multidimensional poverty

One well-known normative motivation to measure multidimensional poverty arises because poor people's lives can be battered by multiple deprivations that are each of independent importance (Sen, 1992). The other key motivation arises from the empirical mismatch between poverty measured in any single space such as income, and additional important single and multidimensional measures of disadvantage. If it were the case that income (or any other unidimensional measure) were a sufficiently good proxy of other disadvantages for practical purposes (such as targeting or tracking change over time or guiding policy) then, in the interests of parsimony, one might not need to go further. [15]

But empirically, many studies note that the extent of mismatch between key social and income indicators, and even between income and key material deprivations, can be considerable across countries and across groups. For example, Brandolini and D'Alessio (2009) used Italian SHIW 1995 data for six dimensions, and found that the correlation coefficients "show low degrees of association", and that the cross-classifications show "low redundancy". [16] They argue that "*the implied shift towards multidimensionality may certainly originate on purely empirical grounds as being driven by the necessity to enrich the information set and to overcome the deficiencies of monetary indicators*". Similarly Franco and Ruggieri-Laderchi cross-tabulated data in India and Peru on child and adult deprivations in health and education with income poverty, and found that the percentage of people who were capability poor but not income poor, or vice versa, ranged from 21 to 93 percent. [17]

But even if there are discrepancies between individual indicators, it could be that income, being a general-purpose means, is an accurate representative of multiple deprivations. Again, empirical studies have not necessarily substantiated this. Klasen (2000) found that while correlations between expenditure and levels of deprivation in South Africa were strong overall, they were weaker for the most deprived and for certain population groups (Africans, rural, female-headed households, etc.). In that study, 17 percent of those identified as functionings-deprived were not expenditure poor. Other studies focus on certain population groups such as the disabled, and argue that income poverty measures need to be supplemented by information on additional disadvantages. In a 16-country study Mitra *et al.* (2011) find that disability is not significantly associated with consumption pov-

[15] This issue is discussed in Foster and Sen's Appendix 7 of Sen (1997), which discusses various forms of income poverty measures as well as indicators of other functionings.

[16] The need to look beyond correlations is well known and empirically important. To give just one example of many, Jones and Klenow (2010) find a correlation of 0.95 between GDP and their welfare index, but also find that "*across 134 countries, the typical deviation [between the two indices] is around 46%*".

[17] Franco in Stewart *et al,* 2004. See also Klasen, 2000; Qizilbash, 2002; Ruggeri-Laderchi, Saith and Stewart, 2003; Ruggeri Laderchi, 1997; and Ruggieri-Laderchi, 2008.

erty in most countries, but is significantly associated with multidimensional poverty (using different functional forms and thresholds for multidimensional poverty measures) (see also Kuklys, 2005; Zaidi and Burchardt, 2005). And in the European context, Nolan and Marx conclude that the multidimensionality of poverty generally requires multiple variables: "*Both national and cross-country studies suggest that the numbers experiencing high levels of deprivation across a number of dimensions are often quite modest and that low income alone is not enough to predict who is experiencing different types of deprivation: poor housing, neighborhood deprivation, poor health and access to health services, and low education are clearly related to low income but are distinct aspects of social exclusion.*" [18]

Figure 6 *Distribution across combined income poverty and deprivation persistence variable by country*

Countries	Neither persistently income poor nor deprived	Persistently income poor only	Persistently deprived only	Persistently income poor and deprived
Denmark	82.8	6.9	8.9	1.4
The Netherlands	78.8	7.1	7.3	6.8
Belgium	73.0	9.3	8.8	8.9
France	70.8	11.6	8.5	9.0
Ireland	64.8	11.4	9.7	14.0
Italy	68.8	9.2	11.3	10.7
Greece	68.8	11.2	9.9	10.1
Spain	72.7	9.2	8.7	9.4
Portugal	64.5	12.0	11.3	12.2
All	70.7	10.4	9.2	9.7

Source: Whelan Layte and Maitre, 2004.

Other analyses explore the relationships between income poverty and other deprivations across time. For example Whelan, Layte and Maître (2004) study material asset deprivation and income poverty across five waves of the ECHP data in nine European countries, and find no strong direct or lagged relationship between them. Figure 6 above from their paper shows that on average, 70.7 percent of people were neither persistently income poor nor persistently deprived, and 9.7 percent of people were persistently both income poor

[18] Nolan and Marx, 2009. See also Balestrino, 1996; Balestrino and Sciclone, 2001; Brandolini and D'Alessio, 1998; and Chiappero-Martinetti, 2000.

and deprived. However the measures disagreed for 19.6 percent of people. These people were either persistently income poor but not materially deprived (10.4%) or persistently materially deprived but not income poor (9.2%). To use either measure alone would be to overlook half of those deprived in the other. To compare persistent deprivation (18.9%) and persistent income poverty (19.6%) measures individually, one would not know this divergence and might even presume that the reference populations coincided. Also, if each measure were used singly, one would lose information on which households were deprived in *both* ways and which in only one. Such studies of the empirical mismatches between income poverty and other deprivations motivate the by now well-established practice of considering multiple dimensions of poverty.[19]

Considering multiple dimensions does not, however, require a multidimensional poverty *index*. The following section moves to consider why an index might add value. In breve, the key value-added of a rigorously implemented multidimensional poverty index is that it conveys additional information not captured in single-dimensional measures on the joint distribution of disadvantage and the composition of poverty among different multiply deprived groups. It also provides a consistent account of the overall change in multidimensional poverty across time and space. To argue this is not to suggest that single-dimensional measures be abandoned; it is to suggest that they be supplemented.

2.1.1. Measurement approaches to multidimensional poverty

As has been often cited, Sen's 1976 paper "*Poverty: An Ordinal Approach to Measurement*" opens with the following sentence: "*In the measurement of poverty two distinct problems must be faced, viz., (i) identifying the poor among the total population, and (ii) constructing an index of poverty using the available information on the poor.*" [20] Based on that paper, most poverty measurement methodologies include the two components of identification and aggregation. Whereas in income poverty measures, a person is identified as poor if their income falls beneath a poverty line, identification in multidimensional space is more complex because it may involve the identification of deprivations with respect to each dimension as well as across dimensions.[21]

Poverty measures that employ data on multiple dimensions can be broadly distinguished according to which of the following operations they include, and the order in which these are conducted. While the details vary, broadly speaking four steps can be identified (Figure 7):

(i) apply dimensional cutoff(s) to identify whether a person is deprived in a dimension ;[22]
(ii) aggregate across dimensions;
(iii) identify whether each person is multidimensionally poor;
(iv) aggregate across people.

[19] Building upon this, recently a multidimensional index was adopted at the European level, which combines income with material deprivation and unemployment data to provide a more accurate assessment of economic deprivations. See for example http://ec.europa.eu/eu2020/pdf/115346.pdf at page 12.

[20] Sen, 1976. In that paper, Sen focuses on aggregation, because the recent literature at that time had focused on identification.

[21] These issues are discussed extensively in Alkire and Foster, 2011*b*; this paper draws upon that account.

[22] This discussion refers to the "person" as a unit of analysis for ease of presentation. Similar measures could be constructed for distinct units of analysis such as the household or some population subgroup like youth or women, or an institution such as a school or health clinic.

The first methodology in Figure 7 engages the same component steps as income or consumption poverty measures. When the component variables can be meaningfully aggregated, a cutoff can be set across the aggregate attainments to identify who is poor, and a poverty index constructed in the same way as for unidimensional poverty. This is depicted in the first column below.

Figure 7 *Order of operations*

	1 Unidimensional	2 Multidimensional (Marginal)	3 Multidimensional (Joint)
Apply Deprivation Cutoffs	n/a	1	1
Aggregate Across Dimensions	1	3	(2)
Identify Who is Poor	2	n/a	3
Aggregate across People	3	2	4

Source: author.

Bourguignon and Chakravarty confine the term "multidimensional poverty" to measures that use cutoffs for each dimension or attribute: "*the issue of multidimensionality of poverty arises because individuals, social observers or policy makers want to define a poverty limit on each individual attribute: income, health, education, etc...*"[23] Hence the other two methods in Figure 7 might be categorized as multidimensional because they apply deprivation cutoffs to multiple dimensions; however, only the last approach necessarily identifies whether each person is multidimensionally poor.

The marginal approach (column 2) uses deprivation cutoffs to identify who is deprived in a particular dimension. It then aggregates information across a population to generate a deprivation *measure* for each dimension. The vectors of marginal deprivation measures are then aggregated. Note that people are identified as deprived or non-deprived with respect to each dimension individually; the measure does not identify people as "multidimensionally" poor or non-poor. Nor does it reflect the joint distribution of deprivations. We refer to such indices as marginal (Alkire and Foster, 2011b; see also Anand and Sen, 1997; Atkinson, 2003; Jenkins and Micklewright, 2007).

Marginal indices, being insensitive to joint deprivation, do not require all variables to come from the same survey. Also, they can directly aggregate deprivations that pertain to different reference populations – such as children and adults, or rural and urban popu-

[23] 2003, p.25; emphasis in the original. Other approaches such as the "counting" approaches widely implemented in Europe and the Unsatisfied Basic Needs approaches in Latin America also use this approach (Atkinson, 2003; Feres and Mancero, 2001).

lations. They can also bring together deprivations that occur with different frequencies or orders of magnitude in the population. For example, a marginal measure might combine an indicator on the percentage of people living in households without access to sanitation (a relatively frequent event let us suppose, with data for all households) with an indicator on maternal mortality per 100,000 women, and an indicator on child malnutrition for children under the age of three. Yet consider a man whose house has sanitation but whose wife perished in childbirth and whose young child is malnourished. Is he poor? Marginal measures do not identify each *person* in the society as multidimensionally poor or non-poor so could not answer this question. They might not fulfill the identification criterion of Sen (1976), which might require that identification clarify whether each person in the population was poor or non-poor.

The third column above provides the general order of aggregation for multidimensional poverty measures that reflect joint distribution. As the empirical examples in the previous section demonstrated, in many situations, information regarding simultaneous deprivations might contribute independent value to an overall assessment of poverty. Measures reflecting joint distribution first apply a set of deprivation cutoffs in order to identify the dimensions in which that person is deprived. These measures then identify whether each person is multidimensionally poor.

For example, a person is identified as multidimensionally poor by the *union* approach if the person is deprived in any dimension (Atkinson, 2003; see also Duclos, Sahn and Younger, 2006). A person is identified as multidimensionally poor by the *intersection* approach if and only if she or he is deprived in *all* dimensions. In both of these cases, identification is accomplished by considering the vector of deprivations, but aggregation across dimensions is not required. Alternative identification methods – such as our dual cutoff approach[24] – may require aggregation across dimensions. The multidimensional poverty measure aggregates across poor people to construct an overall measure of multidimensional poverty for the society.

A key point to note is that the joint approach alone identifies people as being multidimensionally poor on the basis of their joint or simultaneous deprivations. This methodology requires all data to be available for each person, which in many cases means that the data must originate from the same household survey. This methodology also requires a common unit of analysis. If the unit of analysis is the person, then each person may be identified as poor based on their own direct deprivations. But the unit of analysis might also be the household, or youth aged 15-24, or it might be a school or health clinic.

The distinctions between these broad approaches to multidimensional poverty measurement are vital and are often overlooked, creating considerable confusion.The remainder of this paper focuses on multidimensional measures that reflect the joint distribution of disadvantage, in order to probe

[24] James Foster and I have adopted an intermediary approach, in which a person can be identified as multidimensionally poor if they are poor in some (weighted) sum or "count" of dimensions that can include union and intersection as well as intermediary cutoffs (Alkire and Foster, 2007).

more completely their characteristics and value-added.

The relevance of understanding interconnections among multiple deprivations was highlighted in the 2009 Report of the Commission on the Measurement of Economic Performance and Social Progress, which argues that "*Some of the most important policy questions involved 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*". The report also highlights the particular relevance of joint distribution when studying disadvantage: "*For example, the loss of quality of life due to being both poor and sick far exceeds the sum of the two separate effects, implying that governments may need to target their interventions more specifically at those who cumulate these disadvantages*" (Stiglitz *et al.* 2009:55). The conclusion affects both survey design as well as the development of summary measures: *"Developing measures of these cumulative effects requires information on the 'joint distribution' of the most salient features of quality of life across everyone in a country through dedicated surveys."* [25]

The identification and aggregation steps of one multidimensional measure reflecting joint distribution will now be illustrated in the following sections, using our most basic and applicable index M_0 drawn from the class M_α introduced in Alkire and Foster (2007; 2011*a*).

2.1.2. One multidimensional poverty measure: the M_0

This section briefly introduces the Alkire-Foster (AF) class of M_α measures that build on the FGT index. We describe our general measurement approach thus:

"A methodology M for measuring multidimensional poverty is made up of an identification method and an aggregate measure (Sen, 1976). Following Bourguignon and Chakravarty (2003), we represent the former using an identification function $p: \mathbb{R}^d \times \mathbb{R}^d_{++} \rightarrow \{0,1\}$, which maps from person *i*'s achievement vector $y_i \in \mathbb{R}^d$ and cutoff vector z in $\mathbb{R}^d_{++}$ to an indicator variable in such a way that $\rho(y_i; z) = 1$ if person *i* is poor and $\rho(y_i; z) = 0$ if person *i* is not poor. Applying ρ to each individual achievement vector in y yields the set $Z \subseteq \{1,\ldots,n\}$ of persons who are poor in y given z. The aggregation step then takes ρ as given and associates with the matrix y and the cutoff vector z an overall level $M(y; z)$ of multidimensional poverty. The resulting functional relationship $M: Y \times \mathbb{R}^d_{++} \rightarrow \mathbb{R}$ is called an *index*, or *measure*, of multidimensional poverty. A methodology is then given by $\mathbf{M} = (\rho, M)$" (Alkire and Foster, 2011*a*, p.477).

Let us consider poverty in d dimensions across a population of n individuals. Let $y = [y_{ij}]$ denote the $n \times d$ matrix of achievements for i persons across j dimensions. The typical entry in the achievement $y_{ij} \geq 0$ represents individual i's achievement in dimension j. Each row vector $y_i = (y_{i1}, y_{i2}, \ldots, y_{id})$ gives individual i's achievements in each dimension, whereas each column vector $y_{\cdot j} = (y_{1j}, y_{2j}, \ldots, y_{nj})$ gives the distribution of achievements in dimension j

[25] Consideration of these issues can be traced to Atkinson and Bourguignon (1982).

across individuals. To weight the dimensions, define a weighting vector w whose j^{th} element w_j represents the weight that is applied to dimension j. We set $\sum_{j=1}^{d} w_j = d$, that is, the dimensional weights sum to the total number of dimensions.

The M_0 measurement methodology can be summarized as follows. Let $z_j > 0$ be the deprivation cutoff in dimension j, and z be the vector of deprivation cutoffs. Define a matrix of deprivations $g^0 = [g^0_{ij}]$, whose typical element is defined by $g^0_{ij} = w_j$ when $y_{ij} < z_j$, and $g^0_{ij} = 0$ when $y_{ij} \geq z_j$. From the matrix construct a column vector c of *deprivation intensity*, whose i^{th} entry $c_i \sum_{j=1}^{d} g^0_{ij}$ represents the sum of the entries in a given row, and represents the weighted deprivations suffered by person i.

Next, identify who is multidimensionally poor. Select a poverty cutoff k, such that $0 < k \leq d$ and apply it across this column vector c. A person is identified as poor if his or her weighted deprivation count $c \geq k$. This can be called *a dual cutoff* identification method, because it uses the *deprivation* cutoffs z_j to determine whether a person is deprived or not in each dimension, and the *poverty* cutoff k to determine who is to be considered multidimensionally poor.

Construct a second matrix $g^0(k)$, obtained from g^0 by replacing its i^{th} row g^0_i with a vector of zeros whenever $c\,i < k$. This matrix contains the weighted deprivations of exactly those persons who have been identified as poor and *excludes* deprivations of the non-poor. M_0 is the mean of the matrix $g^0(k)$. That is $M_0 = \mu(g^0(k))$, where μ denotes the arithmetic mean operator.

M_0 can also be expressed as the product of the (multidimensional) headcount ratio (H) and the average deprivation share among the poor (A). H is simply the proportion of people that are poor, or q/n where q is the number of poor people. A is the average fraction of deprivations poor people experience – $A = \sum_{i=1}^{n} c_i(k)/dq$ – and reflects the average *intensity* of multidimensional poverty.

M_0 satisfies dimensional monotonicity: if a poor person becomes deprived in an additional dimension, the M_0 will increase. M_0 is also decomposable by population subgroups. Additionally, after identification, M_0 can be broken down by dimension. The intuitiveness of M_0 – that it is the product of headcount (H) and intensity (A) – together with these three properties in particular, enable M_0 to be taken apart in various ways to generate multiple insights, as the next section details.[26]

If data are cardinal and satisfy additional assumptions, we identify other measures within the M_α family that can be computed to reflect the depth and severity of multidimensional poverty. These replace the binary g^0 matrix with a matrix of normalized gaps, or with normalized gaps raised to a positive power α; apply the identification function, censor deprivations of the non-poor, and take the mean of the respective matrices to generate other measures.

[26] M_0 also satisfies other properties: replication invariance, symmetry, poverty focus, deprivation focus, weak monotonicity, non-triviality, normalisation, and weak re-arrangement (Alkire and Foster, 2011*a*). These axioms are joint restrictions on the identification and aggregation methodologies.

2.1.3. Unfolding M_0

A relevant feature of the AF methodology is that M_0 (and other measures in that class) can be unfolded directly into multiple meaningful indices, which clarify the extent and composition of poverty in a coherent way. This "dashboard" of internally consistent measures provides a fuller account of the information summarized in the overall poverty measure. The key partial and intuitive indices for M_0 are the multidimensional poverty headcount (H) and the measure of intensity (A). M_0 can also be broken down by dimension, to generate censored headcounts for each dimension, and dimensional contributions to poverty. M_0, H and A can be decomposed by population subgroups such as regions or ethnic groups. Intensity (A) can be broken down by levels to explore who is poorest or least poor. These indicators are used alongside the summary measure to provide a depth of understanding and policy insight that is not possible from the overall measure alone. All of these numbers are generated from the censored matrix $g^0(k)$, and are briefly introduced below.

Partial Indices

1. Headcount (*H*). The percentage of people who are identified as multidimensionally poor. In multidimensional as in unidimensional poverty, the headcount is familiar, intuitive and easy to communicate. It can be compared directly with an income poverty headcount, or with the incidence of deprivations in another indicator, and also compared across time.

2. Intensity (*A*). The percentage of weighted dimensions in which the average poor person is deprived. Intensity reflects the extent of simultaneous deprivations poor people experience. Its lower bound is the percentage k/d (the poverty cutoff as a percentage of the total number of dimensions) and its upper bound is 100%.

Breakdown by Dimension (post-identification)

3. Censored Dimensional Headcounts H_j. From the censored matrix, the mean of each column generates the percentage of people who are both multidimensionally poor and deprived in each dimension.

4. Percentage Contribution of each Dimension to Multidimensional Poverty. The intensity figure *A* can be broken down by dimension to show the percentage that each (weighted) dimension contributes to poverty. This is similar to the property of factor decomposability (Chakravarty *et al.*, 1998), but uses the censored matrix $g_0(k)$.

Decomposition by Subgroup

5. One can decompose the M_0, H, and A by population sub-group to show how each of these varies by region, by ethnicity, by rural and urban areas, or other subgroups for which the sample is representative. These could be used to create poverty maps, for example.

Breakdown by Intensity

6. The intensity (A) is constructed as the mean of each person or household's deprivation count c_i (with appropriate sampling weights applied). The average hides inequality in intensities across people. So the intensity can be broken down into different bands, to show the percentage of poor people who experience different levels of intensity or to target the poorest of the poor.

Related Indices (from g^0 rather than $g^0(k)$ matrix)

7. Raw Dimensional Headcounts. From the raw matrix g_0 the mean of the column vector

generates the "raw" headcounts of persons who are deprived – whether or not they are multidimensionally poor. This can be compared with the censored headcounts to see, for example, which deprivations are most common among the non-poor, or the percentage of deprived persons who are also multidimensionally poor. This can also be useful to prioritize service delivery and to decide between universal versus targeted delivery mechanisms.

2.2. One particular application of M_0: the MPI

In 2010, the UNDP Human Development Report Office and the Oxford Poverty & Human Development Initiative (OPHI) released an acute Multidimensional Poverty Index (MPI) for 104 developing countries (Alkire and Santos, 2010; UNDP, 2010). This section sets out briefly how the MPI was constructed and provides an overview of its results.

2.2.1. Parameters

The MPI generates multidimensional poverty measures by analysing existing publicly available data sources. In particular, the 2010 MPI is based on Demographic and Health Surveys (DHS) for 48 countries, on Multiple Indicator Cluster Surveys (MICS) for 35 countries, and on the World Health Survey (WHS) for 19 countries. Distinct surveys were used for Mexico and urban Argentina. All surveys used were between 2000 and 2008. All questions for each country were drawn from the same household survey for that country.

The MPI implements the first measure in the dual-cutoff approach of Alkire and Foster (2011*a*) introduced above: M_0. This methodology was chosen because it can be used for ordinal and categorical data; it builds upon the FGT income poverty measures, is straightforward and intuitive in construction, and satisfies a number of desirable axioms. In terms of policy relevance, the resulting measure can be broken down by population group and broken down by factor to show the composition of poverty, hence can describe how the extent and composition of multidimensional poverty varies across states or ethnic communities, or across time.

The MPI is constructed using ten indicators covering three dimensions. The three dimensions [27] are health, education, and standard of living. The indicators are nutrition (anthropometric measures) and child mortality for health; years of schooling and school attendance for education; and electricity, water, sanitation, cooking fuel, flooring and asset ownership for living standard. Each dimension is equally weighted at one-third. Each indicator within a dimension is also equally weighted. Thus the health and education indicators are weighted at one-sixth each, and the standard of living at one-eighteenth.

Deprivation Cutoffs: the MPI first identifies who is deprived in each of the 10 indicators. The indicators, cutoffs and weights are summarized in the figure below. Note at this point that we take the household as the unit of analysis. For standard of living indicators, a person is deprived if their household is deprived in that particular indicator. However

[27] It is important at this point to draw attention to an important change in terms: in the methodological discussion each entry in the matrix constituted a "dimension". In describing the MPI, the terminology changes, and each entry in the matrix is termed an "indicator"; the term dimension is used in the MPI to reflect conceptual categories ("health") that do not appear in the g^0 matrix directly.

for health and education indicators, a person's deprivations depend on the achievements of other household members. We will return to this issue and call for further research on this combination of individual and household level data.

Figure 8 *Dimensions, indicators, cutoffs and weights of the MPI*

Dimension	Indicator	Deprived if …	Weight
Education	*Years of Schooling*	No household member has completed 5 years of schooling	16.67%
	School Attendance	At least one school-aged child is not attending school years 1 to 8	16.67%
Health	*Child Mortality*	A child has died	16.67%
	Nutrition	Any adult or child for whom there is nutritional information is malnourished	16.67%
Standard of Living	*Electricity*	The household has no electricity	5.56%
	Cooking Fuel	The household cooks on wood, dung or charcoal	5.56%
	Floor	The household's floor is dirt, sand or dung	5.56%
	Sanitation	The household does not have adequate sanitation (according to the MDG guidelines) or it is shared	5.56%
	Water	The household does not have clean drinking water (according to the MDG guidelines) or it is more than a 30-minute walk away	5.56%
	Assets	The household does not own more than one of: radio, television, telephone, bicycle, motorbike, or refrigerator, and does not own a car or truck	5.56%

Source: adapted from: Alkire & Santos (2010).

Poverty Cutoff: once it has been identified who is deprived in each indicator, the next step is to determine who is multidimensionally poor. The second cutoff, called the "poverty cutoff "*k*, is set across the weighted sum of a person's deprivations. In the case of the MPI, every person is identified as multidimensionally poor if and only if they are deprived in at least one-third of the weighted indicators.[28] That is, a person is poor if he or she is deprived in any two health or education indicators, in all six standard of living indicators, or in three standard-of-living and one health or education indicator.

[28] When there are ten indicators present, the persons who are identified as poor by k=3 are the same persons who would be identified as poor by k=3.33.

Whenever the poverty cutoff *k* requires deprivation in more than one indicator, there will be people who, despite experiencing some deprivation will not be considered multidimensionally poor, simply because their total weighted deprivations is less than the *k* poverty cutoff. A person might cook with wood, but have a separate kitchen and ventilation system, so that it does not indicate poverty in their case. An uneducated person may nonetheless be a self-made millionaire blossoming with good health. Furthermore, deprivations may also be caused by inaccuracies in the data themselves; or by inappropriate indicators for that context. In some climates and cultures a natural floor may not indicate deprivation, for example. Finally, in some cases the data may be inaccurate or people may voluntarily abstain from some dimension due to personal values: for example they may have a low body mass due to fasting or fashion.

In the AF methodology, those deprivations are censored since they are assumed to correspond to people who are not multidimensionally poor. Their values are replaced by zeros in the $g^0(k)$ matrix (which differs from the g^0 matrix precisely in the censoring of these deprivations). Subsequent *MPI analyses are not based on the original raw data* (that would appear in a poverty profile or dashboard, for example, and that is contained in the g^0 matrix) but rather *on the deprivations of multidimensionally poor people.* This censoring of any deprivations of non-poor people is a novel step, so is easily overlooked. It influences all subsequent analysis, at times considerably as is mentioned below.

The MPI, as the more general M_0 measure, is the mean of the censored matrix of weighted deprivations. It can equivalently be calculated as the product of the headcount or *incidence* of poverty – the percentage of people who are multidimensionally poor – and the *intensity* or average proportion of weighted deprivations a poor person experiences. For example, if a person is deprived in nutrition, years of schooling, and three standard of living indicators, then her intensity is 50 percent (1/6 + 1/6 + 3/18). If – on average – every person in a country is deprived in 50 percent of the weighted indicators, and 40 percent of the population are poor in that country, then the MPI for that country is 0.20.

Data considerations

The indicators which could be compared across the DHS, MICS, and WHS datasets were limited in several ways; indeed data limitations proved to be a binding constraint for the MPI. Alkire and Santos (2010) provide a detailed description of the limitations, which include the use of surveys from different years, the fact that not all indicators were present for all countries, that some households did not have eligible populations, and that some subgroups are systematically excluded from the household surveys. These data limitations affect national accuracy as well as cross-country comparability. The MPI values cannot be used to compare the 104 countries' acute poverty in a definitive way, as they are drawn from different years, vary in the definition of certain variables, and some countries lack indicators. The study does claim to:

a. provide a more comprehensive and accurate baseline of acute multidimensional poverty reflecting joint deprivations than is possible using a dashboard of the same indicators,

b. provide an estimate of acute multidimensional poverty in each of the 104 countries using available information about three core dimensions of human development, and

c. demonstrate the AF methodology for measuring multidimensional poverty – which can be adapted to national or regional settings having different objectives or more and better data.

The 2010 MPI results thus form a baseline for each country, drawing on the most recent publicly available dataset containing the MPI indicators. When it is updated using comparable surveys and sampling frameworks, changes over time can be measured and analyzed.

2.2.2. Illustrative results [29]

About 1.7 billion people in the 104 countries covered–32 percent of the entire population – are poor according to the MPI. As the aim of the MPI [30] is to complement income poverty measures with a direct measure of deprivation, Alkire and Santos (2010) compare the MPI headcount to the income poverty headcounts in those countries that have data for both measures (93 of the 104 countries), and find that it lies between the $1.25 and $2/day poverty. Across these countries, 26 percent of the population are estimated to live on $1.25 a day or less and 49 percent live on less than $2 a day. At the national level, they find a clear overall relationship between income and multidimensional poverty, but considerable differences for particular countries. The MPI captures deprivations in health and educational outcomes and key services such as water, sanitation and electricity – hence the literature would predict some mismatch. The MPI and income values will also differ due to different survey years as well as measurement error and data inaccuracies. Further analysis is required to understand the differences and potential complementarities more fully.

In terms of regional distribution of acute multidimensional poverty, the study finds that 51 percent of the world's poor as measured by the MPI live in South Asia (844 million people) and 28 percent in Africa (458 million). In Sub-Saharan Africa 64.5 percent of people are MPI poor; in South Asia it is 55 percent. The intensity of poverty – the average number of deprivations experienced by each household – is also greatest in Sub-Saharan Africa and South Asia. Multidimensional poverty in both continents is troubling both in regards to the number of people who are multiply deprived and the intensity of their poverty.

The study compares the headcount (H) and its intensity (A) across the 104 countries and finds a disconcerting relationship: countries with a higher incidence of multidimensional poverty tend to have a higher average intensity. The study shows that decompositions reveal considerable disparity in MPI among population subgroups – a finding that stimulated decompositions of over 50 countries' MPI values, and analyses of the spatial disparities that emerge (Alkire Roche Santos and Seth, 2011).

The study also breaks down the MPI by indicators. This is a post-identification decomposition, hence results still exclude the deprivations experienced by those not identified as poor. This decomposition reveals the structure of poverty among the poor. For example, among the Kikuyu ethnic group in Kenya,

[29] This section very briefly mentions some results from Alkire and Santos, 2010; for a fuller discussion the reader is referred to that paper.

[30] In all of these figures, 2007 population data are used; it would also be possible to apply the population data from the year in which the survey was conducted in each country.

deprivation in child mortality and malnutrition (both health indicators) contribute most to the poverty. Deprivations in electricity, sanitation and cooking fuel, contribute most to the poverty of the Embu, another ethnic group. Decomposition of poverty in India and Bolivia also reveals interesting differences among ethnic, caste, and religious groups.

Is the MPI robust to a range of weights?[31] As an initial exploration, Alkire and Santos (2010) estimate the MPI using three additional weighting structures: (i) giving 50 percent weight to health and 25 percent weight each to education and standard of living, (ii) giving 50 percent weight to education and 25 percent weight each to health and standard of living, and finally (iii) giving 50 percent weight to standard of living and 25 percent weight each to health and education. The pairwise country comparisons show that 88 percent of pairwise comparisons of country rankings are robust for all weighting structures. They also find the Kendal Tau-b correlation coefficients between the MPI rankings and each of the three new methods to be above 0.90. Alkire, Santos, Seth and Yalonetzky estimate the concordance between all four rankings using several methods. The concordance is high (0.975 and above) and as expected the null hypothesis of rank independence across the four rankings is rejected with 99 percent confidence by the Friedman test. In terms of large changes in ranking, among the 60 countries whose MPI scores range from 0.05 to 0.64, five countries exhibit rank changes of 10 or more places.[32] On the basis of this, we conclude that the MPI country rankings are quite robust to weights.

Analyses are also underway to explore trends in MPI over time for a number of countries. For example, in Bangladesh, 68 percent of people were MPI poor in 2004; by 2007, multidimensional poverty had fallen to 58 percent. Although progress was made in a number of indicators, an improvement in school attendance was the most striking aspect of poverty reduction in Bangladesh. In contrast, Ethiopia reduced poverty by improving nutrition and water, whereas Ghana improved several indicators at once.[33]

2.2.3. Operationality

The MPI is a very elementary international baseline multidimensional poverty measure that is operational. It is deeply constrained by data, but it does implement a recognizable methodology, perform robustness tests, and invite improvements.

In the 2006 Grusky and Kanbur volume, Bourguignon argues that "*the key challenge in the field of poverty analysis is clear. It consists of building a set of instruments, starting with a satisfactory definition of poverty, that would meet part or all of the critiques of the [income poverty] paradigm described above, while retaining at least part of its 'operationality'. Current economic analysis of poverty clearly falls short of this objective… The poverty income paradigm is presently often used in situations calling for alternative definitions of poverty, essentially because instru-*

[31] The results in this paragraph draw on Alkire, Santos, Seth and Yalonetzky, 2010.
[32] Alkire, Santos, Seth and Yalonetzky, 2010.
[33] Apablaza, Ocampo and Yalonetzky, 2010.

ments to handle these definitions are not available. The challenge is to create those instruments, rather than trying to make the initial paradigm artificially fit a different conceptual basis" (pp.78-79).

Essentially, the most basic claim of the MPI is that it is an operational instrument, whose strengths and limitations have been made quite clear, and which can be developed and strengthened in the future. The remainder of this paper seeks to identify further issues and catalyze research by which to advance.

2.3. Research questions and debates

The above summarizes the MPI, and the methodology that underlies it. The MPI is one possible implementation of M_0, which itself is one measure within the Alkire-Foster (AF) family of $M\alpha$ measures. The $M\alpha$ family is itself one possible approach to multidimensional poverty measurement that reflects joint distribution. The MPI has come under scrutiny and criticism; it could be useful to categorize the "discontents" in terms of their generality. Many of the issues raised with respect to the MPI would be shared by any multidimensional poverty measure that reflects the joint deprivations that poor people experience – for example, Bourguignon and Chakravarty (2003), Bossert *et al.* (2007), or a counting headcount. The latter include *methodological* issues, such as aggregation, weights and cutoffs; *data* issues, such as whether it is possible to obtain sufficiently accurate data on relevant dimensions from one survey; *political* issues, such as updating and manipulation; and *economic* issues, in particular the link between multidimensional poverty measures and welfare economics. Some issues pertain to the AF methodology directly – such as its neutrality with respect to compensation among dimensions, and the focus axiom. Others pertain to the implementation of AF methodology in the Multidimensional Poverty Index (MPI).

This final section discusses four sets of issues: weights, unit of analysis, data, and aggregation into a single index. They are issues in two senses: first, they may be "critiques" that have been offered; and second, they may be questions for future research and innovation. I articulate most issues with respect to the MPI and the M_0 methodology; however many questions would be answered differently by different measurement approaches. The debate thus far has also passed over some issues within these categories that may be of equal or greater importance to those already articulated, so I take the liberty of proposing these as well.

2.3.1. Joint distribution measures

As was signaled above, poverty measures that reflect the joint distribution of deprivations for one person or household proceed by aggregating [weighted] information on deprivations across all dimensions for each person, identifying multidimensionally poor persons on that basis, and subsequently aggregating across poor people to construct a poverty measure. This section identifies some areas for further work, both on general-purpose methodologies and on how these methodologies are implemented in practice.

Weights

Shortly after the release of the MPI, Martin Ravallion (2011*a,b*) drew attention to, among other things, questions regarding the robustness of the MPI to a plausible range of weights

as well as to the space in which weights were articulated.[34] Recall that the MPI sets weights between the incidence of deprivations, with health, education and standard of living being equally weighted, and the weights entering in a linear form.

Amartya Sen among others sees the need to set weights in multidimensional measures as a strength not an embarrassment: "*There is indeed great merit… in having public discussions on the kind of weights that may be used*" (1997). After all, any national budget implicitly sets weights on many dimensions of welfare, often with little debate.The weights on the MPI are explicit: equal weights on each dimension, and on each indicator within a dimension. Yet given the legitimate diversity of human values, Sen also argues that it may not be necessary to agree on a precise set of weights: ideally, measures would be developed that are robust to a range of weights. However it is pertinent to ask why weights are required, and in particular to note that in the absence of weights, in many situations it will not be possible to identify a plausible group of people who are multidimensionally poor. This section outlines the reason weights are needed – both methodologically and also empirically – and summarizes the MPI robustness results, then identifies issues for research.

Weights can enter at the identification and/or aggregation steps; in the AF methodology they often enter in both steps. Observe that if either the union or the intersection identification approaches are used, *no particular weights on dimensions are required in order to identify who is poor*. Recall that the union approach identifies someone who is deprived in any dimension as poor. The intersection approach requires a person to be deprived in all dimensions in order to be identified as multidimensionally poor. The appeal of union or intersection approaches is simplicity: neither requires specific weights to be set across the dimensions. Note however that both do impose some non-zero weight on each dimension. Indeed, the selection of dimensions and cutoffs plays a very significant role in shaping the results, but weights are not required. Why then does the AF methodology introduce a different approach?

A key reason to develop an intermediary approach to identification is practical. In empirical applications that have greater than two dimensions, often both union and identification approaches give quite extreme values. Of course, this depends upon the dimensions, indicators, and cutoffs selected in each context. The intersection approach can be rather stringent. For example, if we were to identify who was poor using the intersection approach for the same indicators and weights used in the Multidimensional Poverty Index (MPI) launched in 2010 (Alkire and Santos, 2010), we find that the average percentage of poor people across 104 countries is 0 percent, and in only one country would more than 2 percent of people be identified as multidimensionally poor, Burundi (5.1%).

[34] It will be interesting and fruitful to relate this discussion to the distinct but related discussion of the "indexing problem". See Arneson, 1990; Arrow, 1973; Fleurbaey, 2007; Rawls, 1971; Sen, 1991*a* and references therein.

Table 4 *Union and intersection and MPI identification using the MPI deprivation matrix, 2010*

Identification Method	Average Headcount (H) [35]	Number of the 104 countries with H > 90%	Number of the 104 countries with H < 90%
Union	58%	32	0
Intersection	0%	0	103
MPI (k=3)	32%	1	30

Source: author.

On the other hand, if we use the union approach, the average (population-weighted) headcount of multidimensionally poor people would be 58.4 percent, with a range from 4.6 percent in Hungary to 99.6 percent in the Central African Republic. In 32 countries more than 90 percent of people would be identified as multidimensionally poor by the union approach. Furthermore, the variation can be quite wide: the difference between the union and the intersection headcounts of multidimensionally poor people is more than 90 percent in 32 of the 104 countries. For example, in Ethiopia, 98 percent of people would be identified as poor according to the union approach, and 0 percent of persons were identified as poor by the intersection approach. So in some contexts, both approaches to identification are rather unwieldy for policy purposes because they identify no one or almost everyone as poor.

If a poverty cutoff to identify the multidimensionally poor is set at any intermediary level between union and intersection, then weights are required. In the AF methodology, general weights are applied to the 0-1 deprivation vectors, where the sum of the weights is equal to the number of dimensions *d*. A person is identified as poor depending upon the sum of the weights of the dimensions in which they are deprived. However other approaches could be implemented. The point to note at the moment is that without weights we cannot identify who is multidimensionally poor *except* by using the union or the intersection method. Weights are needed merely to identify multidimensionally poor persons, even if no summary or aggregate measure is constructed. As is widely understood, all aggregate measures that reflect joint distribution will also impose weights at the aggregation stage on measures. If weights are to be set with transparency and without embarrassment, further research is required on three issues:

Standards and kinds of robustness to weights: first, additional work on robustness to a range of weights is required. The 2010 MPI values might be thought of as a baseline, and are not directly comparable for the data reasons

[35] Population-weighted average of multidimensional poverty headcount using the 2007 population data, the union approach to identification and MPI indicators/cutoffs/weights across all 104 countries. The simple average headcount, with each country equally weighted, is also 58 percent.

already outlined. This limits the power of studies regarding the robustness of country MPI rankings. Furthermore, standards have not yet been established regarding the kinds of robustness to weights a multidimensional poverty measure should satisfy. Also, the AF measures are designed to inform poverty analysis not just rankings by country or subgroup, and the level of multidimensional poverty, as well as its break-down by indicator, is affected by the weighting structure. So methodologies are required which explore the robustness of different relevant descriptive analyses to a range of plausible weights.

Source of Weights: a second key question is how to generate weights, and what the conceptual as well as practical and empirical considerations are in the choice of method. Approaches to setting weights that have been implemented include participatory consultations, survey questions (on time trade-offs, gambling, socially perceived necessities, and subjective well-being), statistical techniques, expert opinion, axiomatic approaches,[36] and, most commonly, normative weights applied by the author.[37] Quantitative as well as conceptual studies have implemented, compared and scrutinized a range of approaches to setting weights, for example in health economics and social policy (Dibben *et al.*, 2007); such studies are required for multidimensional poverty measures.

Space of weights: third, even if the weights on indicators were brilliantly clear, there could be questions about the *space* in which to set normative weights. In the MPI, the weighting vector w applies to the incidence of deprivation. However, one might transform the vector such that the normative weights apply in a different space. For example, if 10 percent of people are deprived in clean water and 40 percent are deprived in sanitation, then there might be arguments for an 80-20 or 20-80 weighting structure. Alternatively, the weights might be fixed, if possible, in the space of public expenditure. The appropriate space will depend on the purpose of the measure as well as the accuracy of potential transformations. It is necessary to set out the alternatives, their implications for the MPI and also the decompositions and further analyses, as well as their potential value-added and limitations.

Unit of Analysis

One of the features of multidimensional poverty measures is that they must construct a row of deprivations for the same unit of analysis. However in many cases household survey data contains information about the household as well as diverse individual members of the household. It would be useful to consider further how to combine individual and household level data when the unit of analysis is the person, and when it is the household.[38]

[36] In the case of Mexico, the legal guidelines governing the development of their national multidimensional poverty measure include principles–for example, that economic and social progress had to balance each other, and that the achievement of a certain level in each social dimension should be seen as a human right. James Foster and I then developed an axiomatic approach showing that these principles, together with assumptions regarding the accuracy of data, were sufficient to set the weights across dimensions and to uniquely identify people as multidimensionally poor. S. Alkire and J.E. Foster, 2009, "Memo to CONEVAL".

[37] See for example Dibben *et al.*, 2007, and Decanq and Lugo, forthcoming.

[38] Of course in other feasible applications of AF measures, the unit of analysis might be an institution (school, community health clinic), a community (datazone), a business or cooperative, or even a state or country (for governance indicators). However in the case of multidimensional poverty the unit is likely to be a person or household.

There are three kinds of combinations that need to be examined in each circumstance: a) how to combine information that is available for each household member (as years of schooling) and how to address "missing values" in some responses or attribute scores to ineligible respondents; b) how to attribute household level data to individuals (taking into account literature and empirical studies on equivalence scales for income and on intra-household inequalities in distribution), and c) when and how it is justified to use a variable from a single respondent or from a subset of household members to represent all household members. These may require detailed expertise on each indicator.

The combination methods must also consider biases due to differently sized households, as well as households with different compositions. In the 2010 MPI, larger households have a greater probability of being deprived in the health and school attendance indicators, and less probability of being deprived in "years of schooling" and at least the "asset" indicator among the standard of living indicators.The overall effect is not clear. Household composition – the age and gender of household members, as well as their relationships to one another – also varies. A household of male migrant workers will have a relatively low probability of being deprived in nutrition (most surveys lack male malnutrition data), as well as child school attendance and child mortality, whereas a household with a great number of children will have a relatively larger probability.

Studies are needed to enumerate alternative methods of combining the data, what errors may be introduced by different methodologies, and how to check the robustness of results to choices made. Empirical studies also are needed to explore the magnitude of differences introduced by different methodologies and to generate examples of careful and rigorously verified methods of combining individual and household data. Alongside quantitative work, qualitative and ethnographic studies can be used to explore the assumptions underlying different alternatives, and consider which equivalence scales and intra-household aggregation methods are most accurate in a given context.[39]

Data

The data restrictions on the MPI or any other global measure that requires internationally comparable indicators are considerable, as was detailed earlier. The data constraints at a country level are less binding, but it can be useful to itemize common constraints. Many of these are well recognized. For example, many household surveys omit institutionalized populations such as the imprisoned, the homeless, and the hospitalized; further, certain surveys exclude key groups such as the elderly or a gender group. The sampling frame, periodicity, and quality of household surveys are also regularly criticized. Multidimensional measures raise a distinctive set of questions in addition to these for two reasons.

Data on each variable must be available for the same person. If a multidimensional poverty measure follows Sen's approach, and identi-

[39] For example, identifying a household as non-deprived if any member has 5 years of schooling, as the MPI, presumes that education is shared across household members; in some cultural contexts or in some kinds of households, that assumption may not be accurate.

fies *who* is multidimensionally poor first, then information on joint deprivations is required. Surveys reflecting joint distribution were advocated by the 2009 Stiglitz-Sen-Fitoussi Commission on the Measurement of Economic Performance and Social Progress. The need for data on different dimensions for the same person or family is a fundamental issue for this class of measures.

In developing countries, all questions may need to come from the same survey (or else be generated for the same household through matching or mapping).[40] Increasingly, multi-topic household surveys have specialized to explore multidimensional health *or* the quality of education *or* empowerment or water management, etc. Such surveys treat one or a few topics in some depth precisely because no single indicator has been identified as a sufficient single proxy for that dimension. A legitimate concern is whether it is possible to construct brief modules on each dimension such that the data generated are sufficiently accurate.[41] This requires the input of professionals from different disciplines and areas of expertise. It also requires participatory and qualitative work to explore the accuracy of indicators and measures after implementation.

Data must be accurate at the individual level. Second, because the data are aggregated first across each person or household, each variable must reflect deprivation at the level of the person or household itself – not merely when averaged.This affects survey design. For example, a question on "morbidity within the past two weeks" may provide useful data on average, but is unlikely to be a good indicator of the respondent's general health status. Related issues arise for indicators of maternal mortality where incidence is very rare. Also, the component indicators of an index must be scrutinized conceptually to identify whether they are stock or flow variables, whether subjective or objective, and whether they refer to resources, inputs, outputs or outcomes. For example the MPI includes child mortality, which could be a stock variable; it was included in part because no better health variables were available, and in part because empirical studies showed that it does change within relevant time periods. Research on each dimension-area is required to propose the shortest and most high quality questions that reflect different dimensions of poverty accurately for the relevant unit of analysis, type of indicator, and time period of interest.

2.3.2. The AF methodology

The AF methodology has certain desirable properties as sketched above, and certainly M_0 is quite applicable due to its ability to use ordinal variables. But there are also a number of issues or questions that arise within the context of that methodology, of which one is discussed below.

Summary index

A clear preliminary question is why to measure multidimensional poverty – in the sense of providing a summary–at all? It is already possible to consider a vector of deprivations. A vector of deprivations can be constructed

[40] When all relevant indicators are not present on the same survey, where feasible, surveys could be matched to provide data for the same individual or household from different surveys. Poverty mapping techniques might also be used, when they provide sufficient accuracy at the individual/household level, and are appropriate for policy formulation.

[41] For example, Browning.

from different data sources; data from the same survey are not required. Also, weights are not required, so a dashboard seems less controversial.

The first part of this question is why to provide an overall summary measure.[42] One foundational reason is that a summary measure defines each person as multidimensionally poor or non-poor, and so is a poverty measure in the sense of Sen (1976); whereas a dashboard or vector of deprivations may define deprivations one by one for different groups but does not look across dimensions at their joint distribution. Also, an overall measure – for example at the national level – allows for comparisons across time (has poverty gone up), and also across regions (which have higher poverty) using a consistent metric.

A further question is why to provide a summary measure and alongside it a vector of deprivations (*e.g.* the censored headcounts). The censored headcounts we provide reflect joint distribution and cannot be generated without the imposition of weights unless union or intersection identification approaches are used. Depending upon the value of the poverty cutoff k and the shape of the distribution, the impact of censoring on single-dimensional headcounts can be quite significant, and noting the discrepancies themselves can be of interest.

To take a first example from the MPI, Table 5 compares "raw" headcounts – that are directly available from the data – and censored headcounts – which reflect our identification step. In the international MPI, deprivation in adequate sanitation in Ghana was 88 percent in the raw matrix, but only one-third of these households were MPI poor, hence in the censored matrix, the percentage of people who are MPI poor and live in households lacking sanitation falls to 30 percent. In Iraq, 29 percent of households had school-aged children not attending school, but only 41 percent of these households were MPI poor, so the censored headcount of people living in MPI poor households where a child is not attending school is 12 percent. In India, 48 percent of people lived in households with at least one malnourished member, but only 80 percent of these people were MPI poor, so the censored matrix shows 38 percent of people are MPI poor and live in households with a malnourished member. In general, across the MPI indicators, the discrepancy between raw and censored matrices was highest in sanitation and cooking fuel, and lowest among years of education.

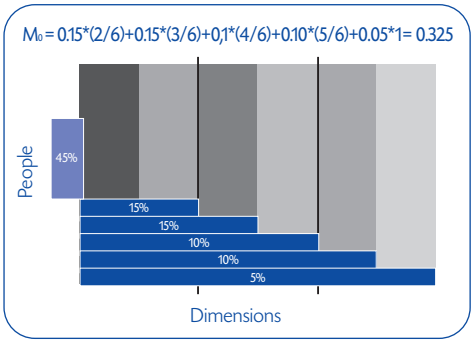
[42] See also Alkire and Foster, 2011*b*.

Table 5 *Selected country differences between raw and censored headcount ratios (percentage decrease)*

Country	School Attend	Years of Education	Child Mortality	Nutrition	Electricity	Sanitation	Water	Floor	Cooking Fuel	Assets
Ghana	15.3%	7.4%	22.8%	17.1%	44.8%	**65.9%**	42.3%	34.0%	65.0%	47.9%
Haiti	5.7%	1.4%	9.9%	13.6%	24.6%	36.7%	25.6%	11.4%	41.0%	22.4%
India	7.7%	3.8%	12.4%	**19.7%**	13.5%	30.8%	24.6%	18.6%	31.0%	22.9%
Iraq	**58.8%**	37.9%	52.6%	50.9%	54.5%	77.3%	70.9%	48.2%	45.4%	55.1%
Lesotho	7.9%	1.7%	11.4%	7.9%	50.4%	47.4%	31.8%	26.0%	40.0%	42.6%
Niger	0.7%	0.3%	1.9%	1.5%	2.6%	5.2%	0.7%	2.2%	7.1%	2.4%
Somalia	3.3%	1.0%	8.3%	1.7%	7.3%	9.5%	7.1%	3.2%	18.7%	7.1%

Source: author.

As the example above illustrates, it is possible and interesting to compare the raw headcounts in each indicator – which includes all deprivations – and the censored headcounts – which focus on people who are deprived in *k* or more weighted indicators.

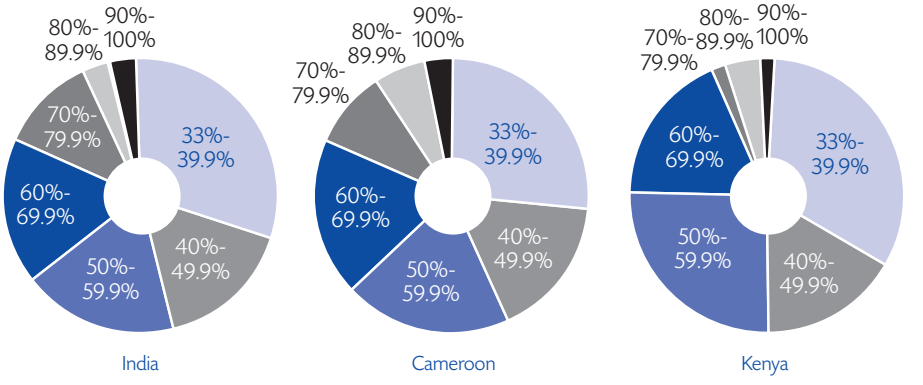


For a second example, we explore further insights from intensity. A distinctive feature of the M_0 measures is the partial index we call intensity. This is constructed, recall, by taking the average proportion of dimensions in which poor people are deprived. For example, in the figure on the left, there are six dimensions and $k = 2$. We can see that 5 percent of people are deprived in 6/6 dimensions (100%), 10 percent each in 4/6 and 5/6 dimensions, and 15 percent each in 2/6 and 3/6 dimensions, so 55 percent of people are deprived in 59 percent of the dimensions on average, where 59 percent is the weighted sum of the above proportions. The area of the blue bars is equivalent to M_0, as would be a single rectangle, 55% x 59%.

Applying the intensity to the headcount creates a measure that can be broken down by dimension: the headcount cannot. Also because the poverty cutoff *k* can vary, one can choose to focus on only a proportion of the population at a time. For example, if we increased *k* to 4, we can see that 25 percent of people would be identified as multidimensionally poor, if to 5, 15 percent, and when $k = 6$, 5 percent of people would be identified as poor. In each successive increment of the value of *k*, the people considered poor are deprived in more dimensions simultaneously.

Because of its construction, analyses of M_0 for any *given* value of k can still describe the composition of *intensity* among poor people – the average proportion of dimensions in which multidimensionally poor people are deprived in a population. In the MPI, a person must be deprived in one-third of the weighted indicators in order to be identified as poor. We could break up the intensity in different ways. In each country briefing for the MPI, we present a pie diagram depicting the percentage of MPI poor people who are deprived in category of intensity. That is, the darkest slice shows the percentage of MPI poor people who are deprived in 90-100 percent of the dimensions (*e.g.* $90 \leq c_i/d \leq 100$). The next lightest represent the percentage of people deprived in 80-89 percent, and so on down to 33-39 percent.[43] As is visually evident the configuration varies between countries. Consider for example India, Cameroon, and Kenya, which are adjacent in the MPI rankings. The average intensity of MPI poverty ranges from 50 to 53.9 percent. However the *composition* of average intensity varies: in Kenya – which has the highest MPI – the percentage of people who are deprived in 30-50 percent of dimensions is just under 70 percent, and about 25 percent are deprived in 60 percent or more dimensions. Cameroon has the highest intensity overall, and the highest percentage of people deprived in 60 to 90 percent of dimensions – about a third.

Figure 9 *Breakdown of intensity of MPI for India, Cameroon and Kenya*



[43] As the MPI weights are .167 and .056 on different dimensions (with 10 indicators), there are additional values of k at which the intensity actually changes; we showed decile bands for clarity only.

Table 6 *The intensity of deprivations among the MPI poor*

India	Cameroon	Kenya
MPI: 0.283	MPI: 0.287	MPI: 0.296
Intensity: 52.7%	Intensity: 53.9%	Intensity: 49.3%

Source: author.

This ability to break apart a multidimensional poverty index by strata of intensity can be useful, for example for targeting. If an institution were able to provide services for 18 percent of the population, and if the multidimensional poverty measure were appropriately constructed for that country, then one could increase the value of the *k* cutoff until it identified about 18 percent of people as poor.[44] These would be the 18 percent of people having the highest intensity of poverty. It can also be analysed further to inform policy interventions: in Cameroon, more people are deprived in 80 to 100 percent of indicators; thus a question for further analysis is who these people are (spatially and in terms of social groups), and whether targeting the poorest poor and providing an integrated spectrum of services might reduce poverty most quickly. Recall that if one brings a high intensity person out of poverty, the reduction in poverty is greater than if one brings a just-poor person out of poverty,[45] because in the former case the average intensity, as well as the headcount, would decrease. It could also be possible to combine an analysis of intensity bands with a subgroup breakdown to show commonalities in the structure of deprivation: for example, if 80 percent of the persons who are deprived in 40-50 percent are deprived in the same two indicators, this suggests different policies than if there was a great dispersion of deprivation combinations. Finally, it is possible and could be useful to undertake these analyses not only for a population as a whole but also for various subgroups – states, ethnic groups, or other subgroups for which the data are representative and the measure is valid.

Hence the summary measure is not a standalone number. Once one has created the censored matrix $g^0(k)$, the mean of which is the MPI, it is natural to generate a range of related descriptions from it: comparisons, subgroup decompositions, analysis by indicator, and by intensity. The data source for these various tools is the post-identification censored matrix $g^0(k)$, created using the achievement matrix, the deprivation cutoffs vector, the weighting vector, and the poverty cutoff. In any identification other than the union approach (in which *k* takes the value of the lowest-weighted indicator), $g^0(k)$ will differ from the deprivation matrix. Analyses based on this matrix are consistent with the summary measure, and serve to unfold its insights. Such analyses can also be compared with the original (raw) data, and with figures from other data sources.

[44] Alkire and Foster, 2009*a*. The degree of precision in adjusting the headcount by increasing *k* depends upon the weighting structure and the number of variables as well as the distribution of deprivations. Note that the poverty cutoff *k* is, in the example given, itself a policy tool (see Alkire and Foster, 2009*a*, section 8).

[45] For example, by identifying exactly the deprivation(s) lacked by those deprived in 30-40 percent in India and Kenya.

2.4. Concluding remarks

This paper has introduced one approach to multidimensional poverty measurement, one particular methodology (AF), one implementation of it (MPI), and four research topics that are either being investigated or are issues for future research. The key strengths of the M_0 methodology are that it is a poverty measure, fulfilling the steps of identification and aggregation that Amartya Sen set out for poverty measures; that it is intuitive and easy to interpret; that it satisfies a set of desirability properties such as subgroup consistency; that it makes explicit the weights set upon dimensions; and that it identifies joint deprivations and has multiple ways of presenting joint deprivation through the measurements such as intensity. Finally, the AF methodology is flexible: the dimensions, cutoffs, and weights can all be chosen to reflect the purpose of the measure and its context; the MPI is only one example out of many possible applications of the underlying methodology.

We used the results from the 2010 104-country Multidimensional Poverty Index (which implemented the M_0 methodology to illustrate some analyses that the measure can generate. We noted that MPI analyses differ from analyses using the original data and indicators because the basic matrix used by all MPI-related figures is "censored" to focus *only* on the disadvantages of people who are jointly deprived in 33 percent (in this case) of dimensions.

The last section briefly introduced an incomplete yet substantive set of research topics, progress in which would take this work to the next stage. These include a set of issues related to many multidimensional measurement approaches – such as work on weights, cutoffs, income, combining individual and household data, policy analysis, linkages to preferences and welfare economics, treatment of ordinal and categorical data, and so on. There are also issues specific to the AF methodology – such as incorporating complementarity and substitutability, and relaxing the focus axiom at identification.

The fundamental question is whether undertaking the further field-building research and collecting missing data is likely to significantly advance various agents' abilities to reduce the incidence, intensity, depth and duration of human poverty. I have argued that investing further in multidimensional poverty measures has the potential to generate significant advances in understanding and to create useful policy tools. To develop this potential, it is vital to establish and convey good practices for the implementation of multidimensional poverty measures, such that measures are implemented with rigour and transparency in the upcoming phase. The paper also identified a few of the many methodological advances that are required. If the methodologies of multidimensional poverty measurement adequately navigate these challenges, they may be seen not as a threat to economics' legitimate parsimony, but as an extension of its core strengths.

References

Alkire, S. and J. Foster (2007), "Counting and multidimensional poverty measures", *OPHI Working Paper series,* 7.

Alkire, S. and J.E. Foster (2009*a*), "Counting and multidimensional poverty measurement", OPHI *Working Paper series*, 32.

Alkire, S. and J.E. Foster (2009*b*), "Memo to CONEVAL", *OPHI Research in Progress.*

Alkire, S. and J.E. Foster (2010), "Designing the Inequality-Adjusted Human Development Index", *OPHI Working Paper series,* 37.

Alkire, S. and J. Foster (2011*a*), "Counting and multidimensional poverty measurement", *Journal of Public Economics*, (95): 476-487.

Alkire, S. and J.E. Foster (2011*b*), "Understandings and misunderstandings of multidimensional poverty measurement", *Journal of Economic Inequality, 9(2): 289-314.*

Alkire, S. and M.E. Santos (2010), "Acute Multidimensional Poverty: A New Index for Developing Countries", *OPHI Working Paper series*, 38.

Alkire, S., M.E. Santos, S. Seth and G. Yalonetzky (2010), "Is the Multidimensional Poverty Index robust to different weights?", *OPHI Research in Progress.*

Alkire, S., J.M. Roche, M.E. Santos and S. Seth (2011), "Multidimensional Poverty Index: An in-depth analysis of the new country results, changes over time and sub-national decompositions", OPHI Research in Progress.

Anderson, G. (2008), "The empirical assessment of multidimensional welfare, inequality and poverty: sample weighted multivariate generalizations of the Kolmogorov-Smirnov two sample tests for stochastic dominance", *Journal of Economic Inequality*, 6(1): 73-87.

Apablaza, M., J.P. Ocampo and G. Yalonetzky (2010), "Decomposing changes in multidimensional poverty for 10 countries", mimeo.

Arneson, R. (1990), "Primary good reconsidered", *Noûs,* 24: 429-454.

Arrow, K.J. (1973), "Some ordinalist-utilitarian notes on Rawls's theory of justice", *Journal of Philosophy*, 70: 245-263.

Asselin, L.-M. (2009), *Analysis of Multidimensional Poverty:* Springer/IDRC, New York.

Atkinson, A.B. (2003), "Multidimensional deprivation: Contrasting social welfare and counting approaches", *Journal of Economic Inequality*, 1(1): 51-65.

Atkinson, A.B. and F. Bourguignon (1982), "The comparison of multi-dimensional distribution of economic status", *The Review of Economic Studies,* 49(2): 183-201.

Balestrino, A. (1996), "A note on functionings: poverty in affluent societies", *Notizie di Politeia,* 12(43/44): 97-105.

Balestrino, A. (1998), "Counting the poor in a fuzzy way: the head-count ratio and the monotonicity and transfer axioms", *Notizie di Politeia,* 14(52): 77-86.

Balestrino, A. and N. Sciclone (2001), "Should we use functionings instead of income to measure well-being? Theory, and some evidence from Italy", *Revista Internazionale di Scienza Soziali,* 109(1): 1-20.

Becker, G.S., T.J. Philipson and R.R. Soares (2005), "The Quantity and quality of life and the evolution of world inequality", *American Economic Review,* 95(1): 277-291.

Bennett, C. and S. Singh, (forthcoming), "Multidimensional poverty: measurement, estimation and inference", *Econometric Reviews,* (47).

Bossert, W., C. D'Ambrosio and V. Peragine (2007), "Deprivation and social exclusion", *Economica,* 74(296): 777-803.

Bourguignon, F. and S.R. Chakravarty (2002), "Multidimensional poverty orderings", *DELTA Working Paper 2002-22,* Paris.

Bourguignon, F. and S.R. Chakravarty (2003), "The measurement of multidimensional poverty", *Journal of Economic Inequality,* 1(1): 25-49.

Bourguignon, F. (2006), "From income to endowments: the difficult task of expanding the income poverty paradigm", in D.B. Grusky and R. Kanbur (eds), *Poverty and inequality,* Stanford University Press, Stanford.

Brandolini, A. and G. D'Alessio (1998), *Measuring well-being in the functioning space,* Banco d'Italia Reseach Department, Rome.

Brandolini, A. and G. D'Alessio (2009), "Measuring Well-Being in the Functioning Space" in E. Chiappero-Martinetti (ed), *Debating Global Society: Reach and Limits of the Capability Approach,* Feltrinelli Foundation, Milan, Italy.

Brighouse, H. and I. Robeyns (2010), *Measuring justice: primary goods and capabilities,* Cambridge University Press, Cambridge, New York.

Browning, M.T. Crossly and G. Weber (2003), "Asking consumption questions in general purpose surveys", *Economic Journa*l, 113(491).

Calvo, C. and S. Dercon (2008), "Risk and Vulnerability to Poverty", in N. Kakwani and J. Silber *The Many Dimensions of Poverty,* Palgrave-MacMillan, Basingstoke.

Calvo, C. (2010), "Vulnerability to multidimensional poverty: Peru, 1998-2002", *World Development,* 36(6): 1011-1020.

Calvo, C. and S. Dercon (2007), "Chronic poverty and all that: that measurement of poverty over time", *Chronic Poverty Reseach Centre, Workshop Concepts and Methods for Analysing Poverty Dynamics and Chronic Poverty, Working Paper Number 89.*

Capppellari, L. and S.P. Jenkins (2007), "Summarizing multiple deprivation indicators", in S.P. Jenkins and J. Micklewright (eds), *Inequality and Poverty Re-examined* (166-184), Oxford University Press, Oxford.

Cerioli, A. and S. Zani (1990), "A fuzzy approach to the measurement of poverty", in C. Dagum, M. Zenga and Università di Pavia (eds), *Income and Wealth Distribution, Inequality and Poverty* (1: 272-284), Springer Verlag, Berlin.

Chakravarty, S., D. Mukherjee and R.R. Ranade (1998), "On the Family of Subgroup and Factor Decomposable Measures of Multidimensional Poverty", *Research on Economic Inequality*, 8:175-194.

Chakravarty, S.R. and D'Ambrosio, C. (2006), "The measurement of social exclusion", *Review of Income and Wealth*, 53(3): 377-398.

Chakravarty, S.R. and J. Silber (2008), "Measuring multidimensional poverty: the axiomatic approach", in N. Kakwani and J. Silber (eds), *Quantitative approaches to multidimensional poverty measurement*, Macmillan, Basingstoke.

Cheli, B. and A. Lemmi (1995), "A 'totally' fuzzy and relative approach to the multidimensional analysis of poverty", *Economic Notes*, 24(1): 115-133.

Chiappero-Martinetti, E. (1994), "A new approach to evaluation of well-being and poverty by fuzzy set theory", *Giornale Degli Economisti e Annali di Economia*, 7-9: 367-388.

Chiappero-Martinetti, E. (2000), "A multidimensional assessment of well-being based on Sen's functioning approach", *Rivista Internazionale di Scienze Sociali*, 108(2): 207-239.

Chiappero-Martinetti, E. (2006), "Capability approach and fuzzy set theory: description, aggregation and inference issues", in A. Lemmi and G. Betti (eds), *Fuzzy Set Approach to Multidimensional Poverty Measurement* (93-113), Springer Verlag, Berlin.

Decanq, K. and M.A. Lugo (forthcoming), "Weights in multidimensional indices of well-being: An overview", *Econometric Reviews.*

Deutsch, J. and J.G. Silber (2005), "Measuring multidimensional poverty: an empirical comparison of various approaches", *The Review of Income and Wealth*, 51(1): 145-174.

Dibben, C., I. Atherton, M. Cox, V. Watson, M. Ryan and M. Sutton (2007). *Investigating the impact of changing the weights that underpin the Index of Multiple Deprivation 2004*, Department for Communities and Local Government, London.

Duclos, J.-Y., D.E. Sahn and S.D. Younger (2006), "Robust multidimensional poverty comparisons", *The Economic Journal*, 116(514): 943-968.

Erikson, R. (1993), "Descriptions of inequality: the Swedish approach to welfare research", in M. Nussbaum and A. Sen (eds), *The Quality of Life,* Clarendon Press, Oxford.

Fleurbaey, M. (2007), "Social choice and the indexing dilemma", *Social Choice & Welfare,* 29: 633-648.

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

Fleurbaey, M. and Gaulier, G. (2009), "International Comparisons of Living Standards by Equivalent Incomes", *Scandinavian Journal of Economics*, 111(3): 597-624.

Foster, J., J. Greer and E. Thorbecke (1984), "A class of decomposable poverty measures" *Econometrica*, 52(3): 761-766.

Foster, J. and A. Shorrocks (1988), "Poverty orderings and welfare dominance", *Social Choice Welfare,* 5(2-3): 179-198.

Foster, J.E. (2007), "A class of chronic poverty measures", *Department of Economics Vanderbilt University, Working Paper No. 07-W01.*

Foster, J.E. and A.K. Sen (1997), "After a quarter century", in A.K. Sen (ed), *On economic inequality*, Clarendon Press, Oxford.

Franco, S. (2004), "Different Concepts of Poverty", in F. Stewart, R. Saith and B. Harriss-White (eds), *Defining Poverty in the Developing World.* Palgrave-Macmillan, Basingstoke.

Gordon, D., S. Nandy, C. Pantazis, S. Pemberton and P. Townsen (2003), *The distribution of child poverty in the developing world*, Centre for International Poverty Research, Bristol.

Grusky, D.B. and R. Kanbur (eds). (2006), *Poverty and inequality,* Stanford University Press, Stanford.

Hirschman, A.O. (1984), "Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse", *American Economic Review*, 74(1): 89-96.

Jenkins, S.P. and J. Micklewright (2007), *Inequality and Poverty Re-examined, Oxford University Press, Oxford.*

Jones, C.I. and P.J. Klenow (2010), "Beyond GDP? Welfare across countries and time" *NBER Working Paper* (16352).

Kakwani, N. and J. Silber (2008*a*), *The Many Dimensions of Poverty*, Palgrave-MacMillan, Basingstoke.

Kakwani, N. and J. Silber (2008*b*), *Quantitative Approaches to Multidimensional Poverty Measurement*, Palgrave-Macmillan, Basingstoke.

Klasen, S. (2000), "Measuring Poverty and Deprivation in South Africa", *Review of Income and Wealth*, 46(1): 33-58.

Kreitler, S. and M.M. Kreitler (2006), "Multidimensional Quality of Life: A New Measure of Quality of Life in Adults", *Social Indicators Research* 76(1): 5-33.

Krishnakumar, J. (2004), *Going Beyond Functionings to Capabilities: An Econometric Model to Explain and Estimate Capabilities*, Cahiers du département d'économétrie, Faculté des sciences économiques et sociales, Université de Genève.

Krishnakumar, J. and P. Ballon (2008), "Estimating basic capabilities: a structural equation model applied to Bolivia", *World Development*, 36(6): 992-1010.

Lelli, S. (2001), *Factor analysis vs. fuzzy sets theory: assessing the influence of different techniques on Sen's functioning approach*, Center for Economic Studies, Leuven.

Lemmi, A. and G. Betti (eds) (2006), *Fuzzy Set Approach to Multidimensional Poverty Measurement:* Springer.

Maasoumi, E. and M.A. Lugo, (2008), The information basis of multivariate poverty assessments, in N. Kakwani and J. Silber (eds), *Quantitative Approaches to Multidimensional Poverty Measurement* (1-29), Palgrave-MacMillan, New York.

McGillivray, M. (2007), Towards a measure of non-economic well-being, in I. Gough and J. A. McGregor (eds), *Wellbeing in developing countries: from theory to research* (143-145), Cambridge University Press, Cambridge.

Mitra, S., A. Posarac and B. Vick (2011), Disability and Poverty in Developing Countries: A Snapshot from the World Health Survey, *World Bank Social Policy Discussion Paper* 1109.

Nolan, B. and I. Marx (2009), "Economic Inequality, Poverty and Social Exclusion", in W. Salverda, B. Nolan and T. M. Smeeding (eds), *Oxford Handbook of Economic Inequality* (315-341), Oxford University Press, Oxford.

Nolan, B. and C.T. Whelan (1996), *Resources, deprivation and poverty*, Oxford University Press, Oxford.

Qizilbash, M. (2002), "A note on the measurement of poverty and vulnerability in the South African context", *Journal of International Development*, 14(6): 757-772.

Ravallion, M. (1996), "Issues in measuring and modelling poverty", *The Economic Journal, 106*(438): 1328-1343.

Ravallion, M. (2011*a*), "Mashup Indices of Development", *World Bank Research Observer*.

Ravallion, M. (2011*b*), On multidimensional indices of poverty, *Journal of Economic Inequality*, 9(2): 235-248.

Rawls, J. (1971), *A theory of justice*, Belknap Press of Harvard University Press, Cambridge, Mass.

Robeyns, I.A.M. and R.J. Van der Veen (2007), "Sustainable quality of life: conceptual analysis for a policy-relevant empirical specification", Netherlands Environmental Assessment Agency (MNP).

Ruggeri Laderchi, C. (1997), "Poverty and its Many Dimensions: The Role of Income as an Indicator", *Oxford Development Studies*, 25(3): 345-360.

Ruggeri Laderchi, C. (2008), "Do concepts matter? An empirical investigation of the differences between capability and a monetary assessment of poverty", in F. Comim, M. Qizilbash and S. Alkire (eds), *The Capability Approach, (203-241)* CUP, Cambridge.

Schokkaert, E. and L. Van Ootegem (1990), "Sen's concept of the Living Standard applied to the Belgian Unemployed", *Recherches Economiques de Louvain*, 56: 429-450.

Sen, A.K. (1976), "Poverty: An Ordinal Approach to Measurement", *Econometrica*, 44(2): 219-231.

Sen, A.K. (1979), "Personal Utilities and Public Judgements: Or What's Wrong with Welfare Economics?", *Economic Journal,* 89(355): 537-558.

Sen, A.K. (1982), *Choice, welfare and measurement*, Basil Blackwell, Oxford.

Sen, A.K. (1987*a*), *On Ethics and Economics*, Basil Blackwell, New York and Oxford.

Sen, A.K. (1987*b*), The Standard of Living: The Tanner Lectures. In G. Hawthorn (ed.), (xiv, 125), Cambridge University Press, Cambridge, New York, Melbourne.

Sen, A.K. (1991*a*), *On indexing primary goods and capabilities*, mimeo, Harvard University.

Sen, A.K. (1991*b*), "Welfare, Preference and Freedom", *Journal of Econometrics*, 50(1-2): 15-29.

Sen, A.K. (1992), *Inequality Re-examined*, Russell Sage Foundation, New York.

Sen, A.K. (1993), "Capability and well-being", in M. Nussbaum and A. Sen (eds), *The Quality of Life*, Clarendon Press, Oxford.

Sen, A.K. (1996*a*), "Interpersonal comparisons of welfare", in A. P. Hamlin (ed.), *Ethics and Economics* (1: 274-292), Elgar, Cheltenham, UK.

Sen, A.K. (1996*b*), "On the foundations of welfare economics: Utility, capability and practical reason", in F. Farina, F. Hahn and S. Vannucci (eds), *Ethics, Rationality, and Economic Behaviour* (50-65), Oxford University Press, Oxford.

Sen, A.K. (1997), *On economic inequality* (Second ed.)., Oxford University Press, Clarendon Press, Oxford, New York.

Sen, A.K. (2002*a*), "The possibility of social choice", in A. K. Sen (ed.), *Rationality and Freedom* (65-118), Harvard University Press, Cambridge and London.

Sen, A.K. (2002*b*), *Rationality and freedom*, Harvard University Press, Cambridge and London.

Stewart, F, C. Ruggeri-Laderchi and R. Saith (2003), "Does it matter that we do not agree on the definition of poverty? A comparison of four approaches", *Oxford Development Studies*, 31: 3.

Stiglitz, J.E., Sen, A., and J.-P. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, http://www.stiglitz-senfitoussi.fr/documents/rapport_anglais.pdf.

Subramanian, S. (2007), "Multi-Dimensional Poverty Measurement with Ordinal Information", *mimeo*, Madras Institute of Development Studies, Chennai.

Thorbecke, E. (2008), "Multidimensional Poverty: Conceptual and Measurement Issues", in N. Kakwani and J. Silber (eds), *The Many Dimensions of Poverty*, Palgrave-Macmillan, New York.

Tsui, K. (2002), "Multidimensional poverty indices", *Social Choice & Welfare*, 19(1): 69-93.

Whelan, C.T., R. Layte and B. Maître (2004), "Understanding the mismatch between income poverty and deprivation: a dynamic comparative analysis", *European Sociological Review*, 20(4): 287-302.

Yalonetzky, G. (2010*a*), Stochastic dominance conditions and test for ordered, discrete variables, *OPHI Research in Progress.*

Yalonetzky, G. (2010*b*), A note on the standard errors of the members of the Alkire-Foster family and its components, mimeo.

Yalonetzky, G. (2011), Conditions for the most robust poverty comparisons using the Alkire-Foster family of measures, *OPHI Working Paper 44.*

Comment: Multidimensional Poverty and its Discontents

Alemayehu Seyoum Taffesse,
International Food Policy Research Institute

The measurement of poverty and comparisons thereof have been an active area of research and praxis, particularly during the last two decades. A strand of the literature thus evolved focuses on the construction of multidimensional measures. The Multidimensional Poverty Index (MPI) of Alkire and Santos (2010) is a recent addition.[46] The MPI achieved a very high profile quickly due to its association with the latest Human Development Report of the UNDP. This note comments on Alkire (2010), which not only presents the MPI to this conference but also addresses some of broader issues relevant to poverty measurement.[47] The comments are based on the version of the paper received on 25th November 2010, and without much reference to companion papers.

(1) Structure and objectives

Alkire (2010) has three main substantive sections. The first outlines the Alkire-Foster multidimensional measures of poverty focusing in particular on M_0. The next section considers the MPI as an application of the M0 measure and reports some illustrative results. The third, and by far the longest single section, explores the way forward in part motivated by comments made on the MPI.

The paper unambiguously states its objective as follows: "*This paper will focus on the question of when, how and why certain multidimensional poverty measures may add value, sketch the limits of the contribution, and introduce a set of standing questions.*" (Alkire, October 2010, p.2).

Readers are thus encouraged to look for these issues in subsequent pages of the paper. In light of the recent discourse on multidimensional measures, the value-added originating in such measures is of particular interest.

(2) Multidimensional poverty measurement: the M_0 and the MPI

The paper begins its discussion of measurement by noting the multidimensionality of poverty – a widely accepted feature. Alkire (2010) justifies the need of multidimensional measures at different levels.

[46] Alkire and Santos (2010): Alkire, S. and M.E. Santos (2010), "Acute Multidimensional Poverty: A New Index for Developing Countries", *OPHI Working Paper Series, 38.*

[47] Alkire (2010): Alkire, S. (October 2010), "Multidimensional Poverty and its Discontents", a paper presented at the EUDN Conference, Paris, 1st December, 2010.

• First, the quality of life is determined, and thus best described, by its various relevant aspects jointly. In other words, what matters is *"the 'joint distribution' of the most salient features of quality of life"* (Alkire, 2010).

• Second, there are observable disparities in outcomes (or assessments) across single dimensions as well as between single and multiple dimensions of well-being. To quote: "*The motivation to take a multidimensional view of poverty arises primarily from the empirical mismatch between poverty measured in any single space such as income and additional important single and multidimensional measures of disadvantage. If it were the case that income (or any other unidimensional measure) were a sufficient proxy of other disadvantages for practical purposes (such as targeting or tracking change over time or guiding policy) then, in the interests of parsimony, one need not go further. A set of indicators–or a multidimensional measure–may add value if no single dimensional index can be found that is adequate for practical purposes.*" (Alkire, 2010, p.4)

• Third, the paper acknowledges that the two reasons mentioned above do not necessarily favour the use of a multidimensional poverty index. A collection of indicators considered individually can in principle do the job. Nevertheless, it proposes that "[*the*] *multidimensional poverty index adds value insofar as it is* ***robust*** *and* ***rigorously implemented,*** *and* ***can convey additional information not captured in unidimensional measures,*** *particularly the joint distribution of disadvantage, and the composition of poverty among the multiply deprived.*" (emphasis added).

With these rationales in the background, Alkire (2010) introduces the class of measures proposed by Alkire and Foster (2007; 2009*a*) and describes a recent application of an element of this class of measures – the Multidimensional Poverty Index (MPI).[48]

The MPI has been developed by the Oxford Poverty and Human Development Initiative (OPHI) and the UNDP. The 2010 Human Development Report contains the MPI for 104 developing countries. The key features of the MPI are:

Dimensions and indicators

The MPI aims to capture three dimensions – health, education, and standard of living – using 10 indicators. The indicators are:

• Health-nutrition (anthropometric measures) and child mortality

• Education: years of schooling and school attendance

• Standard of living: electricity, water, sanitation, cooking fuel, flooring, and asset ownership

Weights

• Each dimension is equally weighted at one-third.

• Each indicator within a dimension is also equally weighted. Thus the health and education indicators are weighted at one-sixth each, and the standard of living at one-eighteenth.[49]

[48] A brief discussion, with an extensive reference list, is included to help locate the proposed index in the broader literature.

[49] Alkire (2010) mentions that *"In fact, as there are 10 indicators, the weights must sum to d = 10, so in fact each dimension weighs 3.33 and k = 3; however for simplicity of presentation we discuss the fractions."* It may be useful to clarify why weights summing to 1 are not used in the methodology section. Greater familiarity may enhance understanding unless there is technical usefulness obtained from using weights summing to number of dimensions.

Data

Already existing surveys provide the data for the construction of the MPI. These surveys are diverse in nature and timing: "*Demographic and Health Surveys (DHS) for 48 countries… Multiple Indicator Cluster Surveys (MICS) for 35 countries, and… the World Health Survey (WHS) for 19 countries. Distinct surveys are used for Mexico and urban Argentina. All surveys used are between 2000 and 2008. All questions for each country were drawn from the same household survey for that country, which enables the MPI to reflect the joint distribution of deprivations a person experiences at the same time.*" (Alkire, 2010, p.7).

Note also that the full set of 10 indicators are available for 63 countries, with the rest missing one or more of the indicators, although all have at least one indicator in all of the three dimensions covered.

Cutoffs

Two types of thresholds are used to identify the multidimensionally poor *via* the MPI.

Deprivation *cutoffs:* The MPI first identifies who is deprived in each of the 10 indicators. A person is identified as deprived or non-deprived in each indicator with respect to a cutoff or threshold selected for each indicator. Some of these cutoffs are linked with the MDGs.

Poverty *cutoff:* In the case of MPI the poverty cutoff is set at 3. Given 10 indicators, a person is identified as multidimensionally poor if and only if he/she is deprived in at least 30 percent of the indicators.

(3) Issues

Objectives

The paper begins with the objective of ascertaining the usefulness of multidimensional measures of poverty over and above unidimensional measures individually or collectively. Two observations in this regard.

It would have been helpful if the discussion about the qualities of the MPI had been complemented by a sufficiently detailed discussion of the shortcomings of unidimensional indicators.

Alkire (2010) identifies "*the empirical mismatch between poverty measured in any single space such as income and additional important single and multidimensional measures of disadvantage*" as key. The significance of this mismatch (the motivating problem) is an empirical question – sufficient systematic empirical evidence (as opposed to the anecdotal and/or the possible) needs to be provided regarding the size of the mismatch and the extent or incidence of practical/policy "errors" induced by such mismatches. For instance, it would have been helpful (and illustrative) to show that the classification of individuals into poor and non-poor will be different when a composite index is used in contrast to each of a collection of specific indicators (that went into the construction of the composite). Comparing a ranking by an indicator with the ranking by the composite would not be sufficient – the two will be different by construction unless the indicator and the composite are so closely correlated that the rankings match one-to-one (a monotonic transformation of each other).

The gains of MPI would then be judged against the costs and challenges of its construction.

Political economy considerations

Some may raise a political economy argument of the kind captured in the following quote from Kanbur (2009):[50] *"In human development, Sen worked with another great economist, the late Mahbub ul Haq, in the formulation of the Human Development Index (HDI). I recall a story that Amartya tells about that development, when he was offering a critique of various technical aspects of the HDI to Mahbub. Paraphrasing Amartya's retelling, Mahbub said something like – I understand the technical issues (which he undoubtedly did – Mahbub was no mean technical economist himself), but my objective is not so much to have the perfect index but rather to have something that can change the terms of the political debate by putting education and health on an equal footing with income. Amartya concludes the story by noting the obvious success of Mahbub's project, a project in which Amartya has of course played a central role."* (p.2)

The message I get from the quote is that that argument has already been won.

Significance

Alkire and Santos (2010, p.65) state: *"The MPI has tremendous practical potential for tracking the MDGs".* The following, taken from Kishtany and Taffesse (2009), makes the broader point of the challenges intrinsic to the MDGs Project.[51]

There is a surprising lack of historical context surrounding the MDGs. Most of the targets are expressed as rates of change in objective variables; two important ones–the achievement of universal primary education and the elimination of educational gender disparities–are in absolute levels. A remarkable feature of the MDGs is that the same quantitative targets are applied to all countries. Many of these countries are, however, at different stages of development. Structural change and sequencing are components of the concept of a stage of development. Countries at different stages have contrasting structural characteristics and patterns of relationships between instruments and objectives; in a dynamic perspective, certain socio-economic outcomes may be prerequisites for the transition to a more advanced stage.

None of this is taken into account in the MDGs' absolute and universal targets. At first sight, the targets seem to be a level playing field for assessing and comparing countries. In fact, that they do not explicitly account for the long term dynamics of structural change tilts the field towards those countries who have built up the critical mass of internal transformations necessary for developmental take-off; countries which have yet to enter this virtuous cycle are at risk of being unjustly chided in the likely event that they fail to achieve many of the goals by 2015.

[50] Kanbur (2009): Ravi Kanbur (July, 2009). "Amartya Sen: A Personal Appreciation," (Comments at a dinner in honour of Amartya Sen at the University of Oxford, June 30, 2009) available at http://kanbur.dyson.cornell.edu/papers.htm.

[51] Kishtany and Taffesse (2009); Kishtany, Naill, and Alemayehu Seyoum Taffesse (April 2009), "Achieving the MDGs – A Note", *Ethiopian Journal of Economics,* 18(1): 101-116.

Taking a longer historical view underlines this point: in the 19th century, when today's rich nations had educational enrolment rates similar to those of today's poor countries, the evolution towards high enrolment was much slower than that seen in many developing countries in recent decades. This is even the case for some developing countries that are on course to "fail" on the educational targets (Clemens, 2004).[52] Today's rich countries only made universal primary education an explicit development goal when they had higher income levels than today's poor countries and had nearly achieved universality. If the MDGs' educational targets had been applied to today's industrialised nations during their own early stages of development, they may well have missed the targets (for similar historical evidence on some of the other targets, see Clemens *et al.*, 2004).[53]

The history of today's rich countries shows that development is a drawn-out, uneven and contradictory process full of reversals and discontinuity. The MDGs, with their ambitious, linear and broad set of socio-economic goals, belie this complexity; contemporary developed countries measured yesterday with today's MDG yardstick might well have been branded "failures".

Issues for further research

The author includes a long list of issues that she identifies for further research: a full third of the paper is devoted to "The way forward: research questions and debates". The paper reflects the laudable desire to acknowledge and deal with issues raised in the rather extensive dialogue that took place after the initial release of the MPI, and certainly tones down the claims made in Alkire and Santos (2010).

While honest and generous, it also makes a reader wonder what has been achieved so far! They also confound the key issue of whether a composite measure of multidimensional poverty (such as the MPI) does better than a collection of indicators deployed carefully and intelligently. For example, some of the problems identified are also shared by unidimensional measures. The issue is whether the totality of problems, both common and unique, seriously limits the validity and usefulness of the multidimensional measure. In my view, while the considerable effort of the author and her team should be applauded, the jury is still out on the matter.

[52] Clemens, M. (2004) "The Long Walk to School: International Education Goals in Historical Perspective", Center for Global Development, Working Paper 37.

[53] M. Clemens, S. Radelet and R. Bhavnani (2004), *Counting Chickens When They Hatch: the Short-Term Effect of Aid on Growth*, Working Paper No. 44, Center for Global Development, Washington, D.C.

Conférences & Séminaires

3. Will GDP Growth Increase Subjective Happiness in Developing Countries?

3. Will GDP Growth Increase Subjective Happiness in Developing Countries?

Andrew E. Clark and Claudia Senik, Paris School of Economics *

Abstract

This paper asks what low income countries can expect from growth in terms of happiness. It interprets the set of available international evidence pertaining to the relationship between income growth and subjective well-being. Consistent with the Easterlin paradox, higher income is always associated with higher happiness scores, except in one case: whether growth in national income yields higher well-being on the long run is still hotly debated. The key question is whether the correlation coefficient is "too small to matter".

The explanations for the small correlation between national income growth and subjective well-being over time appeal to the nature of growth itself (from negative side-effects, such as pollution), and to the psychological importance of relative concerns and adaptation. The available evidence contains two important lessons: income comparisons do seem to affect subjective well-being, even in very poor countries; however, adaptation may be more of a rich-country phenomenon.

Our stand is that the idea that growth will increase happiness in low income countries cannot be rejected on the basis of the available evidence. First, cross-country time-series analyses are based on aggregate measures, which are less reliable than those at the individual level. Second, development is a qualitative process involving take-off points and thresholds. Such regime changes are visible to the eye through the lens of subjective satisfaction measures. The case of transition countries is particularly impressive in this respect: average life satisfaction scores closely mirrored changes in GDP for about the first ten years of the transition process, until the regime became more stable. The greater availability of subjective measures of well-being in low-income countries would greatly help in the measurement and monitoring of the different stages and dimensions of the development process.

* senik@pse.ens.fr, clark@pse.ens.fr. We thank Hélène Blake for precious research assistance and CEPREMAP for financial support. We are very grateful to Oliver Charnoz, Emmanuel Commolet, and Nicolas Gury for incisive comments on a previous draft, and Alpaslan Akay, Alexandru Cojocaru, Luca Corazzini, Armin Falk, Ada Ferrer-i-Carbonell, Carol Graham, Olof Johansson-Stenman, John Knight, Andrew Oswald, Bernard Van Praag, Hillel Rapoport, Mariano Rojas, Russell Smyth, and Oded Stark for advice.

Introduction

Is income growth the only thing that matters in development, and does it raise the level of well-being of the population? *De facto*, economic development is generally identified with growth in GDP *per capita*: international organizations, such as the United Nations Organization, the OECD, the World Bank and the International Monetary Fund, classify countries as developed, intermediate or low-development, depending on whether they are above or below certain thresholds of GDP *per capita*. However, development is of course more than just income growth. It is a multidimensional process, which involves not only a quantitative increase in capital accumulation, production and consumption, but also qualitative social and political changes that enlarge the choice set of the individuals concerned. Institutional progress, human rights, democracy, gender equality and other capacities are an integral part of development. We can then ask whether these qualitative objectives can be attained by maximizing GDP. And in addition, we might worry that income growth will yield negative side-effects, which reduce well-being, such as environmental externalities, the destruction of traditional social links, the concentration of the population in urban and suburban centres, the development of work-related stress, and so on.

"Is Growth Obsolete?" The provocative title of the paper by William Nordhaus and James Tobin (1973) reflects the radical questioning of growth as an engine of well-being. Although the authors answer this question in the negative, many economists and social scientists have come to the conclusion that, in developed countries, economic growth per se has little impact on well-being and should therefore not be the primary goal of economic policy (see Oswald, 1997). How much of this argument can we extend to developing countries? Or should we follow the proposition of Inglehart *et al.* (2008) that material growth, as measured by GDP *per capita*, is welfare-improving in developing countries, as it takes people out of poverty and precarity, but that it is useless in modern and "post-modern" societies where survival is taken for granted and human development becomes the only valuable goal?

This paper will address the relationship between GDP growth and well-being in developing countries through the lens of subjective well-being measures, *i.e.* self-declared satisfaction judgements collected in surveys of nationally-representative samples of the population over the world. Using these measures as a shortcut to people's well-being, we will try to see whether GDP growth is really a proxy for and a valuable route to happiness.

One of the most important but equally most controversial issues in the subjective well-being literature is precisely the income-happiness relationship. In a famous article, Easterlin (1974) ironically asked whether "*raising the incomes of all will raise the happiness of all?*" This was based on the observation that average happiness measures remained flat over the long run in countries which had experienced high rates of GDP growth. The income-happiness nexus has been vividly debated for the past two decades by economists, psychologists and

political scientists. However, most of the evidence to date on the relationship between income and subjective well-being is based on developed countries. Is the Easterlin paradox also valid for developing countries, or is it a rich country phenomenon?

This paper presents an overview of the evidence that has accumulated during the past twenty years of research and illustrates some of the findings using a widely used international database (the World Values Survey, 1981-2005) containing individual life satisfaction and happiness information. In a first section, we present the relationship between income, income growth and subjective well-being and ask to what extent the patterns usually observed in developed countries also hold in developing countries. We discuss the potential existence of a threshold effect in the welfare returns of growth, where the latter are higher in low as opposed to high-income countries. Sections 2 and 3 then present the classic explanations of the Easterlin paradox and their relevance to developing countries. Here, we distinguish the positive and negative side-effects of growth, and the limits to the way in which income can produce subjective well-being that stem from human nature itself (comparison and adaptation effects). Finally, we provide some reasons why we believe that cross-section and panel analysis based on individual data is more reliable than that using aggregated time series. Accordingly, we conclude that the positive income-well-being gradient, supported by individual and cross-sectional data, is difficult to dismiss.

The data used in the paper

This paper essentially hinges on results in the existing literature. However, we have added a number of figures of our own, using the five waves of the well-known World Values Survey (WVS, 1981-2008) database covering 105 countries, including high-income, low income and transition countries, which account for 90% of the world's population. Happiness measures were mostly taken from the WVS and the European Social Survey (ESS): this is the case for 250 out of 368 observations. When happiness data was missing, we used information from the ISSP (101 observations) and 17 observations from the 2002 Latino-barometer. All of these datasets are available at http://worldvaluessurvey.org. The happiness and life satisfaction questions were administered in the same format in all these surveys, with equivalent translations for all countries. The wording of the Happiness question was: "*If you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole?: 1. Not at all happy; 2. Not very happy; 3. Fairly happy; 4. Very happy*". In the WVS, the wording of the Life Satisfaction question was: "*All things considered, how satisfied are you with your life as a whole these days?: 1 (dissatisfied)... 10 (very satisfied)*". The surveys cover representative samples of the population of participating countries, with an average sample size of 1,400 respondents at each wave. We calculated the national average value of the answers to each of these questions (treating them as continuous variables). We also created a misery index defined as the percentage of people who declare themselves to be very happy, or very satisfied, minus the percentage of respondents declaring themselves to be not at all happy, or not at all satisfied. As the results from the two aggregate well-being measures were very similar, we only present here the figures based on average well-being.

The paper also appeals to a measure of trust, which is available in the WVS: "*Generally*

speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?: 1. Most people can be trusted; 0. Can't be too careful". The GDP *per capita* and annual GDP growth information comes from Heston, Summers and Aten – the Penn World Table. We also use other quantitative indicators which are available in the World databank, such as the Gini measure of income inequality, women's fertility rates, adult literacy rates, and life expectancy at birth (see http://data.worldbank.org/). The qualitative indicators of governance were taken from Freedom House and Polity IV (http://www.qog.pol.gu.se/, http://www.freedomhouse.org, http://www.govindicators.org, and http://www.systemicpeace.org/polity/polity4.htm). All these data are available from the World Data Bank: http://www.worldvaluessurvey.org.

Subjective well-being measures: why use them and are they reliable?

The critical quality of subjective well-being is that it is self-reported. Instead of a third person designing some set of criteria (income, health, education, housing, etc.) which will define how well an individual is doing, individuals themselves are asked to provide a summary judgement of the quality of their life. While some have doubted the usefulness of subjective measures, we think that there are fairly compelling reasons to include them in the economists' arsenal.

Think of an individual's level of well-being as being some appropriately-weighted sum of all of the aspects of life that matter to them. There are at least two significant obstacles for it to be measured objectively. The first is that we need to be sure that we cover all of the aspects of life that are important to the individual, and it seems *a priori* difficult to make up a definitive measurable list of these. The second problem is that we have to apply appropriate weights to construct the final well-being index. This might appear problematic right from the start: in the context of the aggregate data used in the Human Development Index, for example, how much is literacy worth in terms of life expectancy? Moreover, it would appear extremely likely that any such weighting will differ between individuals, and probably in ways that it is not easy to observe. It is consequently very tempting to sidestep the difficulties involved by asking individuals to make these calculations themselves, in responding to evaluative questions about their own lives.

The well-being questions asked in this context are often very simple ones, such as "*How dissatisfied or satisfied are you with your life overall?*" (from the British Household Panel Survey), which is answered on a seven-point scale, with one referring to "*Not satisfied at all*", four to "*Neither satisfied nor dissatisfied*" and seven to "*Completely satisfied*". Alternatively individuals may be asked about their happiness, as in the following question from the American General Social Survey (GSS): "*Taken all together, how would you say things are these days, would you say that you are very happy, pretty happy, or not too happy?*" Other questions may refer to positive and negative affect or mental health.

These questions are increasingly widely included in surveys across the social sciences. One reason for their popularity is that they are simple to put into questionnaires, as probably the majority of those that appear are single-item (although there are very many multiple-item scales that are also available in the literature: see http://acqol.deakin.edu.au/

instruments/instrument.php for a summary of some of these). A second point is that the vast majority of respondents seem to understand the question: non-response rates are very low. The third reason, which from our point of view is the most important, is that the answers to these questions do seem to pick up how well people are doing.

This last statement might seem to be rather uncontroversial: after all, we would expect a question on life satisfaction to measure exactly that. The potential problem lies exactly in the subjectivity of the reply. In particular, if individuals understand the question differently, or use the response scales differently, then there is a danger that someone who answers six on a one to seven satisfaction scale is no better off than another person who has given an answer of five. Luckily there is by now a varied body of evidence suggesting that these subjective well-being measures do contain valid information.

A first point to make is that subjective well-being measures are well-behaved, in the sense that many of the correlations make sense. In cross-section data, variables reflecting marriage, divorce, unemployment, birth of first child and so on are typically correlated with individuals' subjective well-being in the expected direction.[54] If the answers to well-being questions were truly random, then no such relationship would be found.

We want to know whether asking A how happy she is will provide information about her unobserved real level of happiness. One simple check, called Cross-Rater Validity, is to ask B whether she thinks A is happy. This work has been carried out in a number of settings (see Sandvik *et al.*, 1993; Diener and Lucas, 1999), including asking friends and family, or the person who administered the interview. Alternatively, we can use individuals who do not know the subject: *B* may be shown a video recording of *A*, or may read a transcription of an open-ended interview with *A*. In all cases, *B*'s evaluation of the respondent's well-being matches well with the respondent's own reply.

Another approach to validation consists in relating well-being scores to various physiological and neurological measures. It has been shown that answers to well-being questions are correlated with facial expressions, such as smiling and frowning, as well as heart rate and blood pressure. The medical literature has shown that well-being scores are correlated with digestive disorders and headaches, coronary heart disease and strokes. Research has also looked at physical measures of brain activity. Particular interest has been shown in the differences in brain wave activity between the left and right prefrontal cortexes, where the former is associated with positive and the latter with negative feelings. These differences can be measured using electrodes on the scalp or scanners. Research has shown (for example, Urry *et al.*, 2004) that these differences in brain activity are correlated with individual well-being responses. These measures of brain asymmetry have been shown to be associated with cortisol and corticotropin releasing hormone (CRH), which regulate the response to stress, and antibody production in response to influenza vaccine (Davidson, 2004). Consistent with subjective well-being and brain asymmetry measuring the same

[54] See, for example, the findings in Di Tella, MacCulloch and Oswald (2003), based on the analysis of the well-being reported by levels of a quarter of a million randomly-sampled Europeans and Americans from the 1970s to the 1990s.

underlying construct, individuals reporting higher life satisfaction scores were less likely to catch a cold when exposed to a cold virus, and recovered faster if they did (Cohen *et al.*, 2003).

The last block of evidence that people "mean what they say" is that, in data following the same individual over a long period of time, those who say that they are dissatisfied with a certain situation are more likely to take observable action to leave it. This phenomenon is apparent in the labor market, where the job satisfaction that the individual reports at a certain point in time is a good predictor of her being observed in the future to have quit her job (examples are Freeman, 1978; Clark *et al.*, 1998; Clark, 2001; Kristensen and Westergaard-Nielsen, 2006). One important subsidiary finding in this literature is that job satisfaction predicts quits even when we take into account the individual's wages and hours of work. This prediction of future behavior seems to work for the unemployed as well as for the employed. Clark (2003) shows that mental stress scores on entering unemployment in the British Household Panel Survey (BHPS) data predict the length of the unemployment spell, with those who suffered the sharpest drop in well-being upon entering unemployment having the shortest spell. This finding has been replicated in using the life satisfaction scores in the German Socio-Economic Panel (GSOEP) data by Clark *et al.* (2010). Outside of the labor market, well-being scores have been shown to predict the length of life (Palmore, 1969; Danner *et al.*, 2001). Satisfaction measures have also recently been shown to predict future marital break-up (Gardner and Oswald, 2006; Guven *et al.*, 2010).

One potential use of the analysis of subjective well-being is that it arguably provides us with information on trade-offs between different aspects of an individual's life. If one extra hour of work per week has the same effect on well-being as does 80 euros in additional earnings per month, then the shadow wage (the wage that would compensate for one extra hour of work) is around 18 euros and 50 cents per hour. Some examples of these well-being trade-offs have appeared in the recent literature. For example, Blanchflower and Oswald (2004, p.1381), using American and British data, came to the conclusion that: "*To compensate men for unemployment, it would take a rise in income at the mean of approximately $60,000 per annum. A lasting marriage is worth $100,000 per annum (when compared to widowhood or separated).*"

This capacity of subjective data to weight the different dimensions of development one against the other (to calculate marginal rates of substitution between two dimensions) is particularly adapted to the multidimensionality of economic development. The structure of the well-being equation, as estimated in a country, can be seen as a synthetic measure that would have aggregated the different arguments of a social welfare function. The usual problem of the social planner (and of the social choice school of normative economics) is indeed to decide on the weights that should be attached to the different arguments of the social objective function. Subjective measures allow this obstacle to be avoided by measuring directly the synthetic result of the weighting alchemy made by individuals themselves. An illustration of this is the paper by Di Tella and MacCulloch (2008, pp.31-33), where the authors use the American GSS and the Eurobarometer to estimate national welfare functions. They propose such marginal rates of substitution:

• Life expectancy/income: "*A person who expects to live one year longer due to the reduction in the risk of death is willing to pay $5,052 in annual income in exchange (6.6% of GDP per capita)*".

• Life expectancy/unemployment: "*In terms of the unemployment rate, denying an individual one year of life expectancy has an equivalent cost to increasing the unemployment rate by 1.1 percentage point*".

• Pollution/GDP: "*a one standard deviation increase in SOx emissions, equal to a rise in 23kg per capita, has a decrease on well-being equivalent to a 1.5% drop in the level of GDP per capita*".

• Inflation/unemployment: "*a 1% point rise in the level of inflation reduces happiness by as much as a 0.3 percentage point increase in the unemployment rate*".

• Crime/GDP: "*a rise in violent crime from 242 to 388 assaults per 100,000 people in the United States (i.e. a 60% rise)... would be equivalent to a drop of approximately 3.5% in GDP* per capita".

• Working hours/GDP: "*a 1% rise in working hours would have to be compensated by a 2.4% rise in GDP* per capita *(to leave happiness unchanged)*".

These examples illustrate the capacity of subjective well-being measures to serve as a useful tool for public policy aimed at maximizing well-being as countries develop.

Before we turn to the evidence on growth and subjective well-being, we should warn the reader of two abusive approximations contained in this paper. First, we use the terms happiness, life satisfaction and well-being indiscriminately. Second, we treat these measures as though they were cardinal, although they are more properly ordinal. In doing so, as do the bulk of economists working on happiness measures, we follow the route opened by Ferrer-i-Carbonnell and Frijters (2004).

3.1. The paradoxical relationship between growth and well-being

One of the main catalysts in the voluminous and rapidly expanding literature on income and happiness has been Easterlin's seminal article (1974; updated in 1995), setting out the "paradox" of substantial real income growth in Western countries over the last fifty years, but without any corresponding rise in reported happiness levels. This finding is paradoxical for a number of reasons. First it runs counter to the popular prior belief that increased material wealth and greater freedom of choice should go hand in hand with higher well-being. In a way, our societies are organized on this implicit principle. Second, it seems to contradict a large body of scientific empirical evidence based on cross-sections of countries, and on within-country individual panel data. This section presents and discusses the available evidence on these contradictory findings, and asks whether the Easterlin paradox is a rich-country phenomenon or also something relevant for policy-makers in developing countries. A summary of the wide-ranging data sources and results appears in the appendix of this chapter.

3.1.1. Income raises happiness in the cross-section

Within-country cross-section

"*As far as I am aware, in every representative national survey ever done, a significant bivariate relationship between happiness and income has been found.*" (Easterlin 2005, p.67).

Almost all of the empirical work based on within-country surveys includes individual income or household income (or more precisely, the log of income) as a control variable to explain well-being. Log income invariably attracts a positive and statistically significant coefficient, of considerable size. It is typically one of the most important correlates of self-declared happiness. "*When we plot average happiness versus average income for clusters of people in a given country at a given time… rich people are in fact a lot happier than poor people. It's actually an astonishingly large difference. There's no one single change you can imagine that would make your life improve on the happiness scale as much as to move from the bottom 5 percent on the income scale to the top 5 percent*" (Frank, 2005, p.67). This holds for both developed and developing countries, even if it has sometimes been suggested that the income-happiness slope is larger in developing or transition than in developed economies (see Clark *et al.*, 2008, for a survey).

Layard *et al.* (2010), for instance, report that within a country, a unit rise in log income raises individual self-declared happiness by 0.6 units on average (on a 10-point scale). Stevenson and Wolfers (2008, p.13) estimate the within-country well-being-income gradient over each of the countries available in a number of international datasets (the American General Social Survey, the World Values Survey, the Gallup World Poll, etc.). They con-

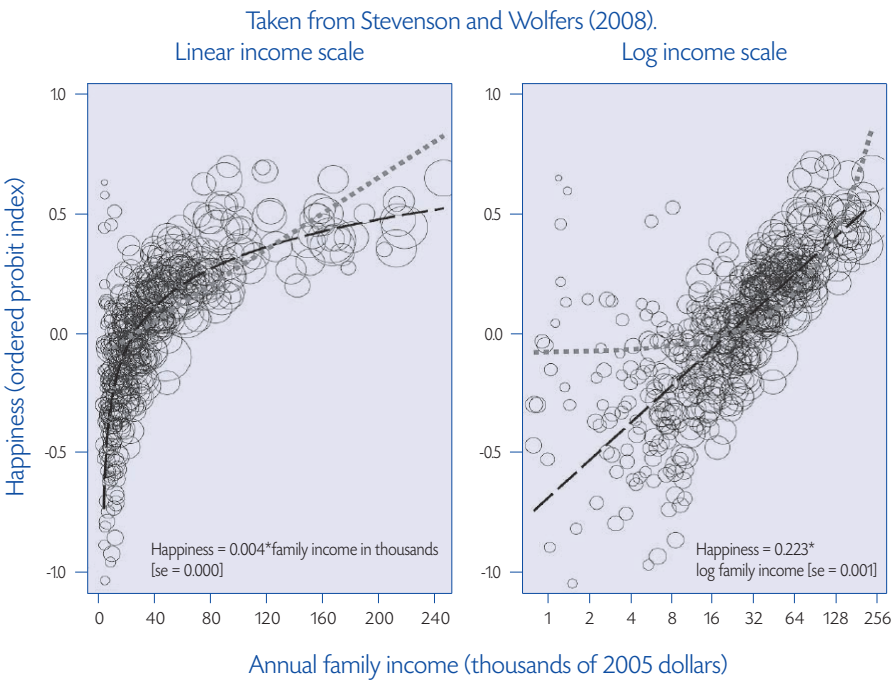
clude that: "*Overall, the average well-being-income gradient is 0.38, with the majority of the estimates between 0.25 and 0.45 and 90 percent are between 0.07 and 0.72. In turn, much of the heterogeneity likely reflects simple sampling variation: the average country-specific standard error is 0.07, and 90 percent of the country-specific regressions have standard errors between 0.04 and 0.11.*"

As an illustration, Figure 10.A depicts the household income-happiness gradient in the United States. The fitted relationship is well described by a log-linear function. The same findings hold in a series of surveys covering developing countries. Figure 10.B shows the income decile-happiness gradient in China in 2007 (based on World Values Survey data): the same positive relationship appears. In general, the fact that in a given society the rich are happier than the poor is a well-established and undisputed empirical finding in this literature.

Figure 10.A

Income and happiness in the American General Social Survey (1972-2006)

Taken from Stevenson and Wolfers (2008).

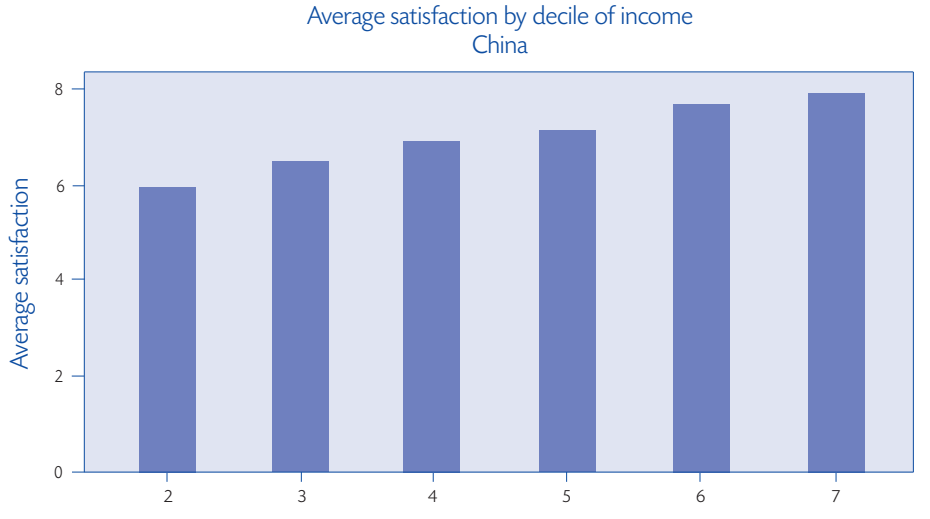


a. Each circle aggregates income and happiness for one GSS income category in one year, and its diameter is proportional to the population of that income category in that year. The vertical axis in each panel plots the coefficients from an ordered probit regression of happiness on family income category x year fixed effects; the horizontal axis plots real family income, deflated by the CPI-U-RS. In each panel the short- and long-dashed lines are fitted from regressions of happiness on family income and the log of family income, respectively, weighting by the number of respondents in each income category x year. Survey question ask. "Taken all together, how would you say things are these days – would you say that you are very happy, pretty happy, or not too happy?"

Source: General Social Survey (USA), 1972-2006; authors' regressions.

Figure 10.B

Income and happiness in a Chinese cross-section.



Source: WVS. China 2007.

Note: *We group together the three deciles (7, 8, 9) which were only rarely reported in the Chinese sample. We have dropped the two extreme deciles.*

Cross-sections of countries

The empirical evidence is even more conclusive and consensual regarding the income-happiness gradient across countries. Deaton (2008), for example, finds an elasticity of 0.84 between log average income and average national satisfaction across a large set of nationally representative samples of individuals living in 129 developed and developing countries, collected by the 2006 Gallup World Poll. In the same spirit, Inglehart (1990, chapter 1) analyses data from 24 countries at different levels of development and finds a 0.67 correlation between GNP *per capita* and life satisfaction. In a more recent paper, Inglehart *et al.* (2008) report a correlation of 0.62 using all available waves of the World Values Survey. Wolfers and Stevenson (2008, p.12), using a very comprehensive set of data, uncover "*a between-country well-being-GDP gradient… typically centered around 0.4*". [55] In the surveys analyzed by Inglehart *et al.* (2008), 52 percent of the Danes indicated that they were very satisfied with their life (with a score of over 8 on a 10-point scale) and 45 percent said they were very happy. On the contrary, in Armenia only 5 percent said they were very satisfied and 6 percent very happy.

Figure 11.A (taken from Inglehart *et al.*, 2008) shows the concave relationship between income *per capita* and average happiness across developed, developing and transition countries of the world, over the 1995-2007

[55] These estimates vary because of the composition of the sample and the controls included in the regressions.

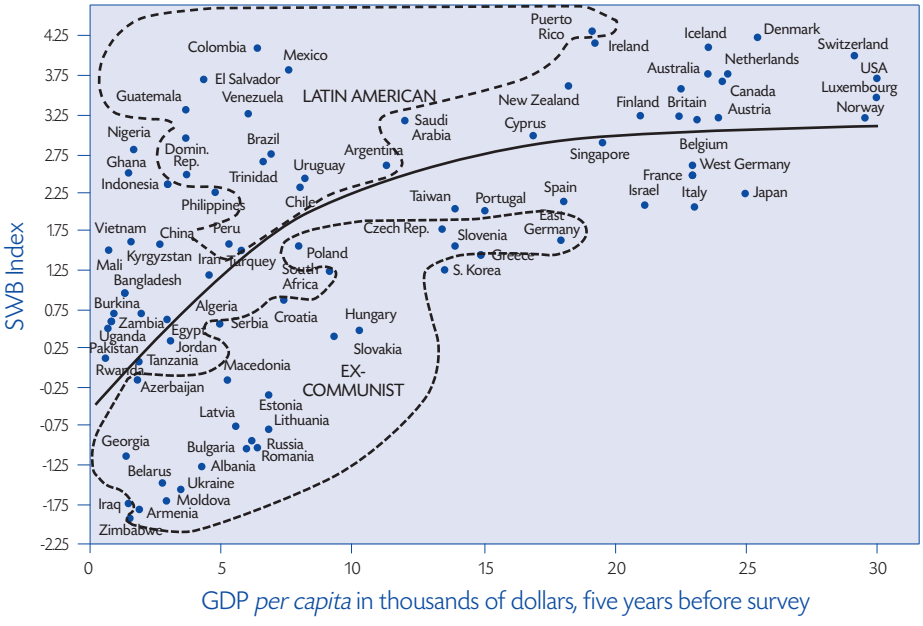
period. A similar graph appears in Deaton (2008) based on the World Values Survey (1996) and the Gallup World Poll (2006), which we reproduce here as Figure 11.B. As shown in Figure 11.C, "*Each doubling of GDP is associated with a constant increase in life satisfaction*" across countries (Deaton, 2008). Figure 11.D illustrates the good fit of a log-linear relationship between income *per capita* and average life satisfaction across countries of the world, in the late 2000s, using the most recent waves of the World Values Survey.

Many other contributions to the "macroeconomics of happiness" have documented the fact that individuals in general report higher happiness and life satisfaction scores in higher-income countries (see for example Blanchflower, 2008), even if certain types of societies seem to be more conducive to happiness than others (Inglehart *et al*, 2008). In Figure 11.A, for example, Latin American countries are systematically found above the regression line, while transition countries form a cluster lying below the regression line tracing out the average relationship in the data.[56]

Figure 11.A

GDP per capita and SWB in the world

Taken from Inglehart, Foa, Peterson, Welzel (2008), p.269



Subjective well-being (SWB), *per capita* gross domectic product (GDP), and different type of societies. Well-being index is based on reported life satisfaction and happiness, using mean results from all available surveys conducted 1995-2007 (cubic curve plotted; *r*=.62). PPP=purchasing power parity estimates.

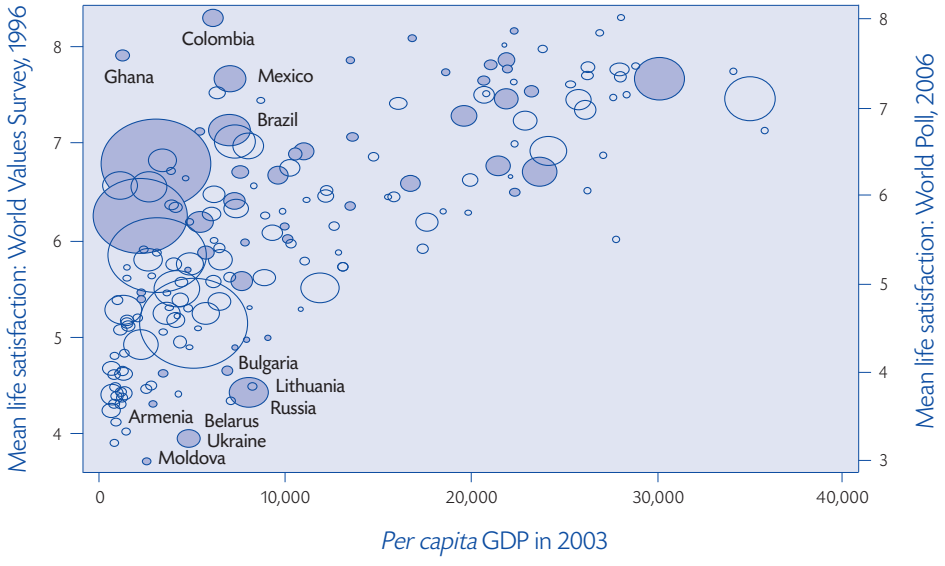
[56] According to Guriev and Zhuravskaya (2009), the reasons for the lower happiness level in transition countries are the deterioration in public goods provision, the increase in macroeconomic volatility and mismatch of human capital of residents educated before transition (unemployment).

Figure 11.B

GDP per capita and Life satisfaction

Taken from Deaton (2008), p.57.

Life Satisfaction in the World Poll and the World Value Surveys
(World Poll data shown as hollow circles, World Values Surveys data as shaded circles)



Source: Penn World Table 6.2.

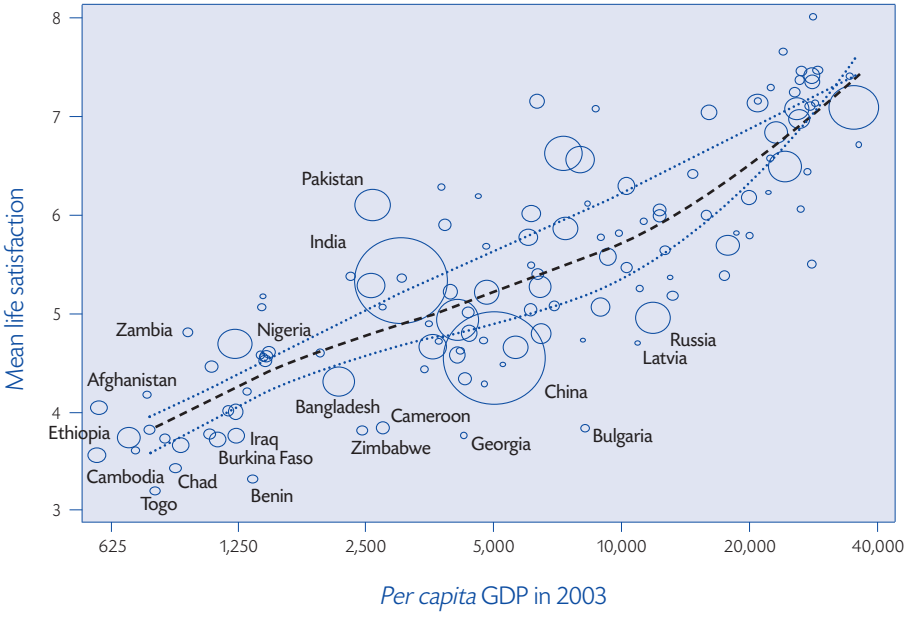
Note: *Each circle is a country, with diameter proportional to population. GDP* per capita *in 2003 is measured in purchasing power parity chained dollars at 2000 prices.*

Figure 11.C

GDP per capita and Life Satisfaction

Taken from Deaton (2008), p.57.

Each Doubling of GDP is Associated with a constant Increase in Life Satisfaction

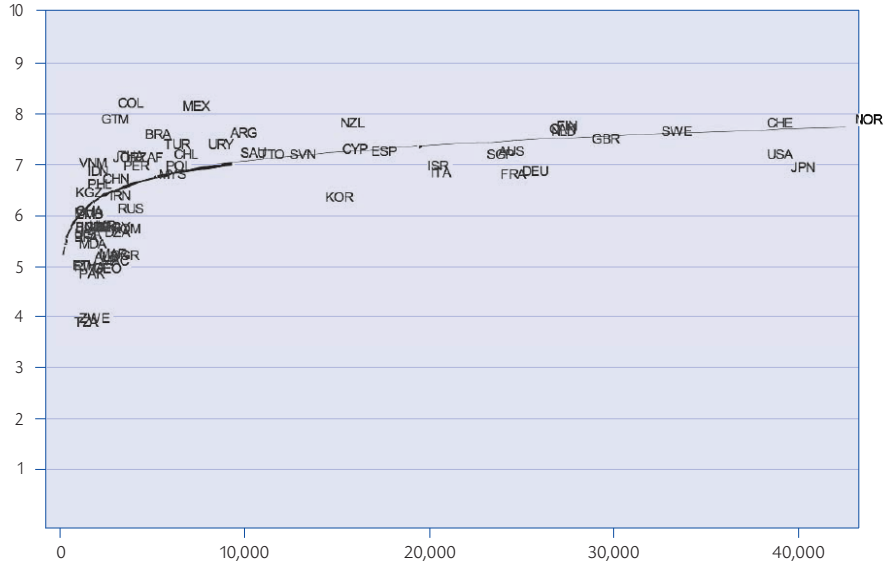


Source: Penn World Table 6.2.

Note: *Each circle is a country, with diameter proportional to population. The scale on the x-axis is logarithmic. The middle line shows average life satisfaction for each level of* per capita *GDP while the outer two lines show the same thing, but for two age groups, ages 15 to 25 – the upper line for most of the figure – and ages 60 and over – which is usually the lower line. GDP* per capita *in 2003 is measured in purchasing power parity chained dollars at 2000 prices.*

Figure 11.D

GDP per capita in the 2000s and Life Satisfaction



Source: WVS.

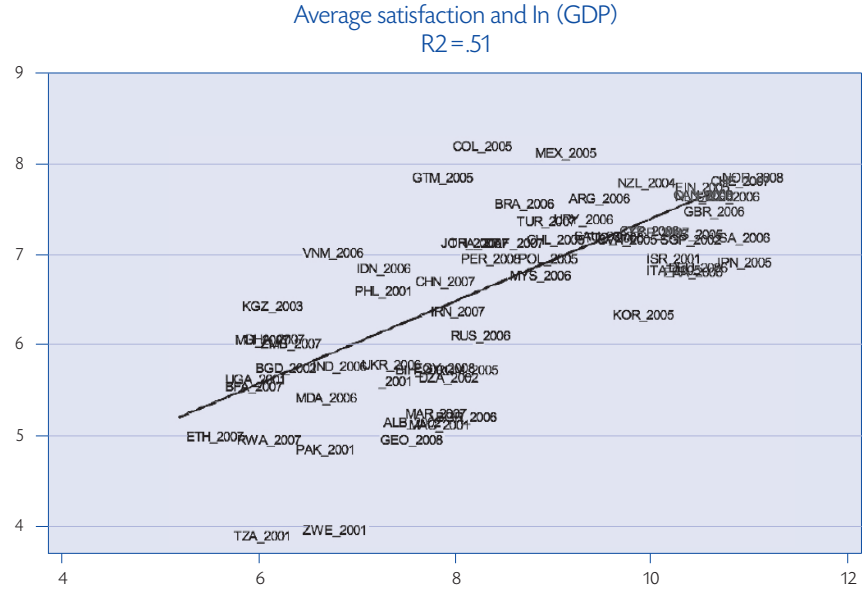
Note: *GDP and average satisfaction are calculated for the last available year for each country (spanning from 2001 to 2008).*

Development and the inequality of subjective well-being

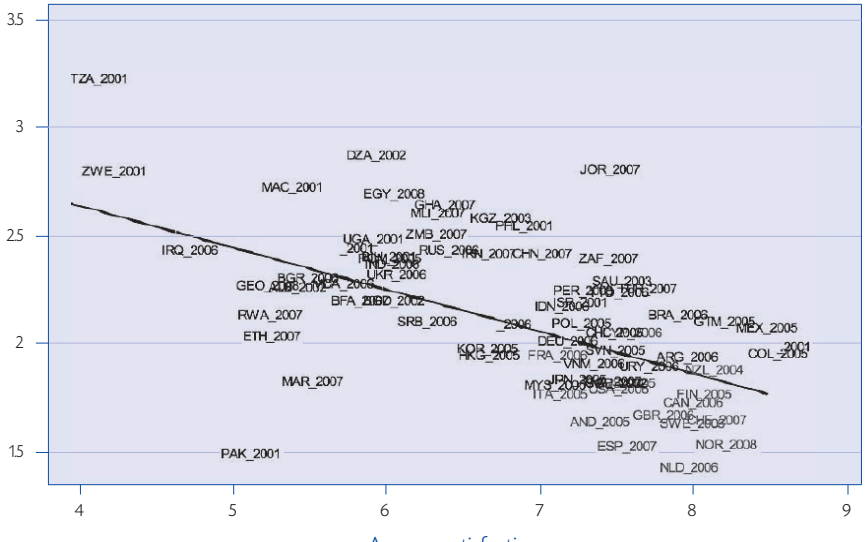
As a complement to the average income-average happiness relationship, we have also looked at the relation between average life satisfaction scores and their standard deviation (treating well-being as a continuous variable). Cross-country analysis produces a striking observation: the higher the average national happiness, the lower the within-country standard deviation of happiness. As such, richer countries have both higher average scores and lower standard deviations of life satisfaction (Figure 12). This suggests one potentially important benefit of GDP growth for low-income countries. If individuals are risk-averse, reducing the variance of SWB in a given society is a valuable objective of public policy.

Figure 12

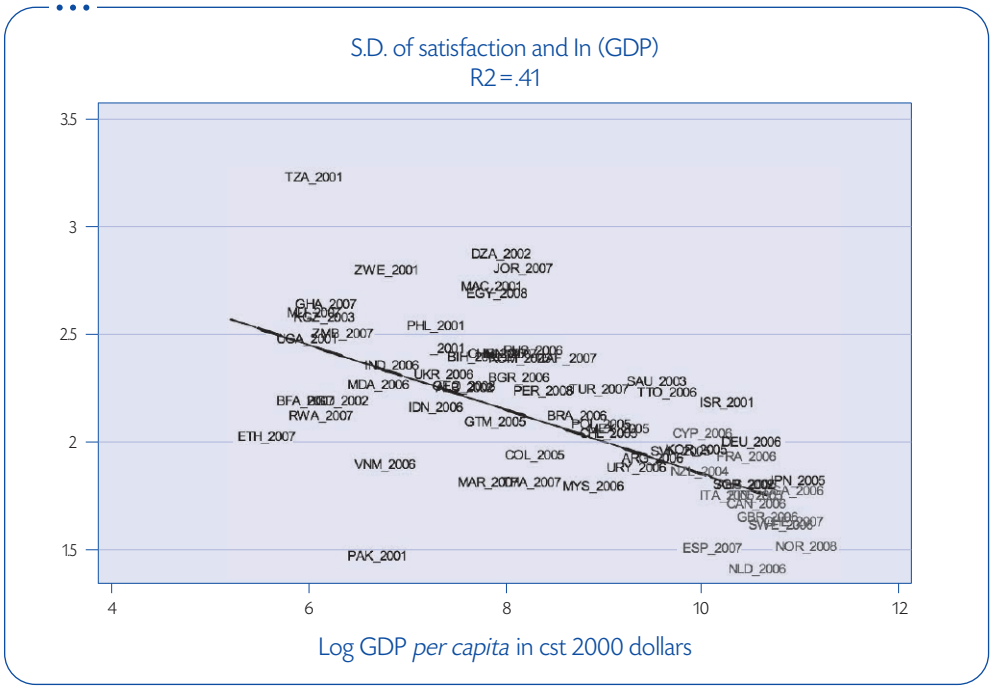
GDP, and the average and standard deviation of happiness



Log GDP *per capita* in cst 2000 dol



Average satisfaction



Source: World Values Survey, 1981-2007.

A positive relation in individual panel data

Thanks to the increased availability of population panel surveys in a number of different countries, a variety of analyses of individual well-being have been able to control for unobserved individual fixed effects, such as personality traits. All of this work has concluded that there is a positive correlation between the change in real income and the change in happiness (see, for example, Winkelmann and Winkelmann, 1998; Ravallion and Lokshin, 2002; Ferrer-i-Carbonell and Frijters, 2004; Senik, 2004 and 2008; Ferrer-i-Carbonell, 2005; Clark *et al.*, 2005). Further, a number of these articles have appealed to exogenous variations in income in order to establish more firmly the causal effect of individual income on happiness (*e.g.* Gardner and Oswald, 2007; Frijters *et al.*, 2004a, 2004*b* and 2006; Pischke, 2010). The slope of the income-happiness relationship is not necessarily the same between groups (Clark *et al.*, 2005; Frijters *et al.*, 2004*a*; Lelkes, 2006). The coefficient on the within-individual change in log income is typically found to be in the vicinity of 0.3 (Layard *et al.*, 2010; Senik, 2005).

There is thus both single-country and international evidence showing that the rich are happier than the poor within a given country, that those in richer countries are on average happier than those in poorer countries, and that an increase in individual income over time is associated with increasing happiness. At this stage then, the evidence is strongly in favour of a development policy based on GDP growth in low-income countries.

3.1.2. The diminishing returns to income growth

The situation is not completely clear-cut, however, as illustrated by the panels of Figures 10 and 11: the positive relationship between income and happiness exhibits diminishing returns. This comes as no surprise to economists, who are accustomed to the idea of the concavity of preferences, *i.e.* decreasing marginal utility and risk-aversion. Concretely, this means that the effect of earning an additional ten thousand dollars on subjective well-being becomes progressively smaller as one's initial level of income increases. This is consistent with the good fit of the log functional form for income-happiness relationship, which is a familiar result in the empirical analysis of subjective well-being across the social sciences.

Is there a threshold in the utility of growth?

"*Once a country has over $15,000 per head, its level of happiness appears to be independent of its income per head*" (Layard, 2003, p.17).

Many authors have suggested a threshold in the welfare effect of income. They recognize that rich countries are happier than poor countries, but believe that there is no strong relationship between GDP *per capita* and happiness among rich countries. This threshold separates "*survival societies*" and "*modern societies*" (Inglehart *et al.*, 2009). It is usually found to be in an interval from $10,000 to $15,000 per annum (Di Tella *et al.*, 2007).[57] Layard (2005, p.149) thus writes: "*if we compare countries, there is no evidence that richer countries are happier than poorer ones – so long as we confine ourselves to countries with incomes over $15,000 per head... At income levels below $15,000 per head things are different ...*". Frey and Stutzer (2002, p.416) similarly claim that "*income provides happiness at low levels of development but once a threshold (around $10,000) is reached, the average income level in a country has little effect on average subjective well-being.*" Even more explicitly, Inglehart (1997, pp.64-65) concludes that "*the transition from a society of starvation to a society of security brings a dramatic increase in subjective well-being. But we find a threshold at which economic growth no longer seems to increase subjective well-being significantly. This may be linked with the fact that, at this level, starvation is no longer a real concern for most people. Survival begins to be taken for granted... At low levels of economic development, even modest economic gains bring a high return in terms of caloric intake, clothing, shelter, medical care and ultimately in life expectancy itself... But once a society has reached a certain threshold of development... one reaches a point at which further economic growth brings only minimal gains in both life expectancy and subjective well-being. There is still a good deal of cross-national variation, but from this point on non-economic aspects of life become increasingly important influences on how long and how well people live...*". The authors continue to reach the same conclusion with updated data: "*Happiness and life satisfaction rise steeply as one moves from subsistence-level poverty to a modest level of economic security and then levels off. Among the rich-*

[57] This notion of a satiation point also goes back to Adam Smith's concept of "*a full complement of riches*", beyond which there could be not be desire for more money. The large landholders of the 18th century had (according to him) reached this limit. However, there may be a limit to the quantity of wealth someone can enjoy in a given society at a certain point of time, but this does not mean that this limit cannot be stretched by the set of new choices brought about by economic growth (*e.g.* the Internet). In other words, the "full complement of riches" could be wider in richer than in less-developed countries.

est societies, further increases in income are only weakly linked with higher levels of SWB" (Inglehart *et al.*, 2008, p.268).

If true, the implication of these findings for developing countries is that GDP growth should be seen as a temporary objective, to be retained only up to a certain level.

But the happiness-log GDP *per capita* gradient does not tend to zero

In spite of these strong claims, the cross-country evidence in favour of such a subsistence level is far from consensual. Bringing together a number of international survey datasets that covering about 90 percent of the world's population, including many developing countries (based on the World Values Survey and the Gallup World Poll), Stevenson and Wolfers (2008, pp.11-12) test for the idea of a cut-point at $15,000 *per capita* per annum (in constant 2000 dollars). They estimate the happiness-GDP *per capita* gradient, and find that: "*the well-being-GDP gradient is about twice as steep for poor countries as for rich countries. That is… a rise in income of $100 is associated with a rise in well-being for poor countries that is about twice as large as for rich countries.*" However, the marginal utility of GDP growth is still positive in developed countries. "*The point estimates are, on average, about three times as large for those countries with incomes above $15,000 compared to those countries with incomes below $15,000. …Taken at face value, the Gallup results suggest that a 1 percent rise in GDP* per capita *would have about three times as large an effect on measured well-being in rich as in poor nations. Of course, a 1 percent rise in U.S. GDP* per capita *is about ten times as large as a 1 percent rise in Jamaican GDP per capita.*"

This result is consistent with Deaton's analysis of the same Gallup World Poll data (Figure 11.B): "*the relationship between log per capita income and life satisfaction is close to linear. The coefficient is 0.838, with a small standard error. A quadratic term in the log of income has a positive coefficient: confirming that the slope is higher in the* ***richer*** *countries! … Using $12,000 of income per capita as a* threshold *between rich and poor countries shows that the slope in the higher income countries is higher! … if there is any evidence for a deviation, it is small and is probably in the direction of the slope being higher among the high-income countries*" (2008). Deaton concludes that *"the slope is steepest in the poorest countries, where the income gains are associated with the largest increases in life satisfaction, but it remains positive and substantial even among the rich countries; it is not true that there is some critical level of GDP per capita above which income has no further effect on life satisfaction"* (2008). In other words, there is indeed diminishing marginal utility to GDP growth, as the level of GDP *per capita* increases, but the return to growth does not converge to zero.[58]

To summarize, an undisputed finding of the happiness literature based on cross-sections of countries is that the relationship between income *per capita* and happiness is concave, *i.e.* has diminishing returns. However, there is no consensus on the existence of a subsistence threshold beyond which the marginal utility of income falls to zero.

[58] It is worth underlying that while the log function is indeed concave, it is not bounded from above. If $y = \log(x)$, then y does not tend to any fixed value as x tends to infinity. Yet, this is the message that a vast majority of specialists in the field have drawn from the decreasing marginal utility of income and the good fit of the log-linear functional form for the relationship between income and happiness.

3.1.3. "Rather than diminishing marginal utility of income, there is a zero marginal utility of income"

The most powerful criticism of pro-growth policy hinges on the empirical evidence regarding the within-country long-run changes in GDP and happiness. Visual evidence provided by Easterlin and his co-authors (1974, 1995, 2005, 2007, 2009 and 2010) illustrates the flatness of the long-run happiness curve plotted against time. One of the most famous and spectacular of these flat curves is shown in Figure 13.A, taken from Easterlin and Angelescu (2007). In spite of the doubling of US GDP *per capita* over a 30-year period (1972-2002), the average happiness of Americans has remained constant. Average happiness is calculated using repeated cross-sections from the American General Social Survey. The same type of pattern has been uncovered in a number of other contributions, with long time-series data covering different developed countries (see Diener and Oishi, 2000). The claim supported by these graphs is radical: in the words of Richard Easterlin," *rather than diminishing marginal utility of income, there is a zero marginal utility of income*" (Easterlin and Angelescu, 2007, p.8).

The absence of any long-run correlation between growth and happiness could be explained by the decreasing marginal utility of income uncovered in the cross-section. However, Easterlin strongly rejects this interpretation: "*The usual constancy of subjective well-being in the face of rising GDP per capita has typically been reconciled with the cross-sectional evidence on the grounds that the time series observations for developed nations correspond to the upper income range of the cross-sectional studies, where happiness changes little or not at all as real income rises.*" But "*the income change over time within the income range used in the point-of-time studies do not generate the change in happiness implied by the cross-sectional pattern*" (Easterlin and Angelescu, 2007, p.24). For example: "*in 1972, the cohort of 1941-1950 had a mean* per capita *income of about $12,000 (expressed in 1994 constant prices). By the year 2000, the cohort's average income had more than doubled, rising to almost $27,000. According to the cross-sectional relation, this increase should have raised the cohort's mean happiness from 2.17 to 2.27. In reality, the actual happiness of the cohort did not change.*"

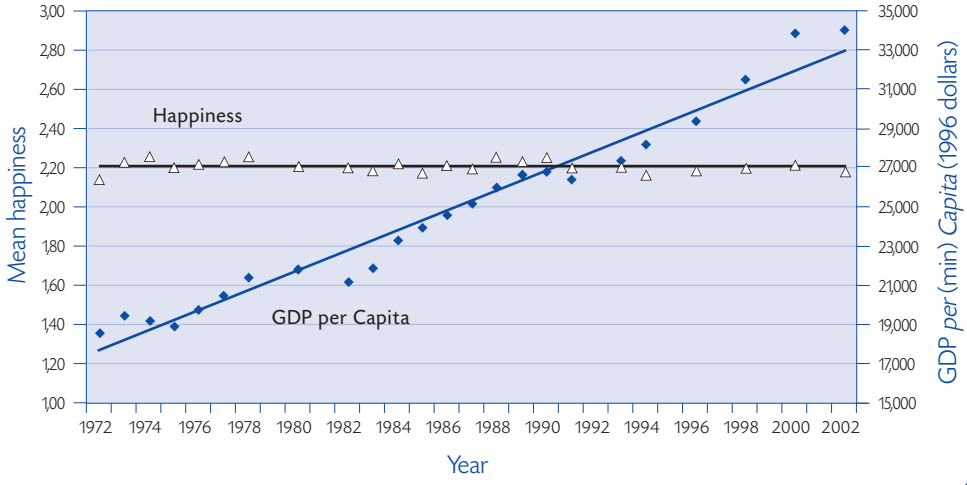
In some of his articles (Easterlin, 2005*a*; Easterlin and Sawangfa, 2005), Easterlin has forcefully underlined that cross-section evidence cannot be transposed to the relationship over time. The change in average self-reported happiness in a country, in the long-run, is not correctly predicted by the instantaneous cross-section relationship between income per head and happiness. Hence: *"knowing the actual change over time in a country's GDP per capita and the multi-country cross-sectional relation of SWB to GDP per capita adds nothing, on average, to one's ability to predict the actual time-series change in SWB in a country"* (Easterlin and Sawangfa, 2009, p.179). This is illustrated in Figure 13.B, taken from Easterlin (2005*a*, p.16), which contrasts the actual (flat) evolution of happiness in Japan, and the predicted (log-linear) change over time.

Hence, the positive concave relationship between GDP *per capita* and SWB, observed in the cross-section, cannot be used to predict the change in SWB in developing countries over time. This new "no bridge" theory underlines

the "fallacy" of transposing cross-sectional relations to time-series data. The lesson for developing countries is that they should not necessarily expect to reach the higher level of well-being that is typical of developed countries by growing over time.

Figure 13.A.

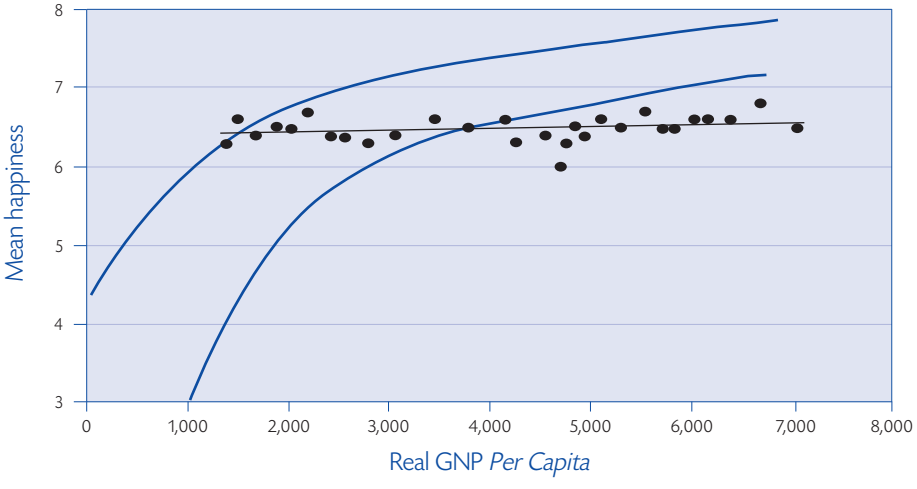
The American paradox. Happiness and real GDP per capita*, United States, 1972-2002*



Source: Easterlin and Angelescu (2007).

Figure 13.B.

Misleading cross-sections. Actual versus *predicted happiness in Japan. 1958-1987*



Source: Easterlin (2005a).

3.1.4. Is the time-series correlation small enough to ignore?

In spite of the spectacular visual evidence offered by Easterlin, his rejection of any correlation over time between growth and happiness is still the object of vivid controversy. In particular, one disputed point is whether the size of the correlation coefficient between SWB and GDP *per capita* is statistically significant, and large. It is small, but is it "*small enough to ignore*"? (Hagerty and Veenhoven, 2000, p.4).

For instance, the absence of correlation between growth and happiness in the fast-developing countries of Japan (after WWII) and China (after 1980) is particularly disappointing. However, Stevenson and Wolfers (2008) have noted a number of discontinuities in the wording of the happiness question and in the sampling of the Japanese cross-sections used by Easterlin. With respect to China, the evidence is scarce (only three points in time) and Hagerty and Veenhoven (2000) underline the fact that the Chinese sample is not representative of the population, as it was initially biased towards more urban demographic groups.

Other work on the long-run macroeconomic time series of happiness has concluded that there is a positive relationship between growth in GDP *per capita* and well-being. Exploiting the World Values Survey, Hagerty and Veenhoven (2003) found that GDP is positively related to the number of "happy life years" in 14 of the 21 countries available in the dataset. In a later paper, Hagerty and Veenhoven (2006) observed a statistically significant rise in happiness in 4 out of 8 high income countries, and 3 out of 4 low income countries.

Inglehart *et al.* (2008) also exploited the most recent waves of the World Values Survey, spanning from 1981 to 2005. They found that, over the complete period, happiness rose in 45 out of the 52 countries for which substantial time-series data is available. Kenny (2005) appeals to data on 21 transition and developed countries and runs regressions of the change in happiness on the growth in GDP, separately for each country. He finds that 88% of correlation coefficients are positive; the overall regression coefficient for all countries together is positive and significant at the 5 percent level.

Inglehart *et al.* (2008) present a series of graphs plotting average happiness against time in different countries, based on the first four waves of the World Values Survey. As they point out: "*in many cases, the results contradict the assumption that, despite economic growth, and other changes, the publics of given societies have not gotten any happier. They show that the American and British series show a downward trend in happiness from 1946 to 1980, but an upward trend thereafter*" [this was confirmed by Easterlin]. "*In general, among the countries for which we have a long-term data, 19 out of the 26 countries show rising happiness levels. In several of these countries – India, Ireland, Mexico, Puerto Rico and South Korea – there are steeply rising trends. The other countries with rising trends are Argentina, Canada, China, Denmark, Finland, France, Italy Japan, Luxembourg, the Netherlands, Poland, South Africa, Spain and Sweden. Three countries (the U.S., Switzerland and Norway) show flat trends from the earliest to the latest survey. Only four countries (Austria, Belgium, the U.K and West Germany) show downward trends*" (the Appendix to Inglehart *et al.*, 2008). Figures 14.A to 14.E taken from

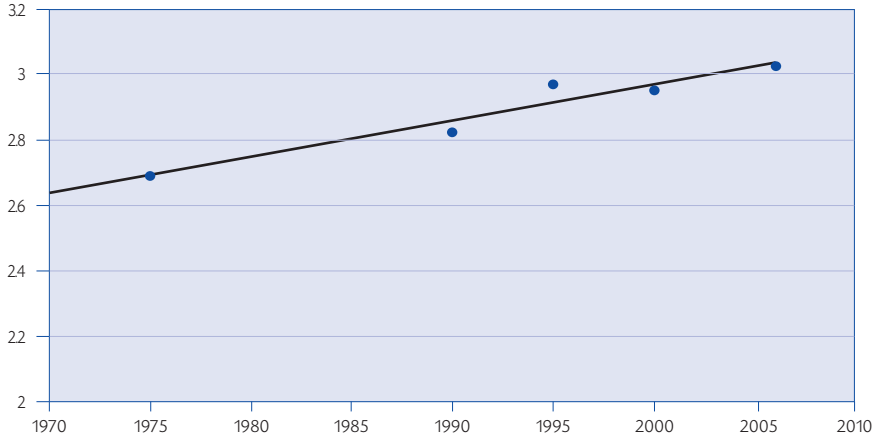
their paper illustrate the positive slope of the happiness curve in India, Mexico, Puerto Rica, South Africa, and the downward slope in China.

Some work has thus uncovered a positive and statistically significant correlation between growth and well-being over time, using within-country time-series data. This includes Hagerty and Veenhoven (2003), Stevenson and Wolfers (2008), and Inglehart *et al.* (2008). In turn, many of these results have been criticized by Easterlin (2005) on the basis of the choice of countries, the confusion between long-run dynamics and the business cycle, and the absence of controls in some of the estimates. Easterlin, with a number of different co-authors, has confirmed and developed his initial conjecture. Authors such as Ed Diener, Rafael Di Tella, Bruno Frey, Robert MacCulloch, Andrew Oswald and Alois Stutzer have provided additional empirical evidence in this direction.

Figures 14.A to 14.E are taken from Inglehart *et al.* (2008, Statistical appendix).

Figure 14.A.

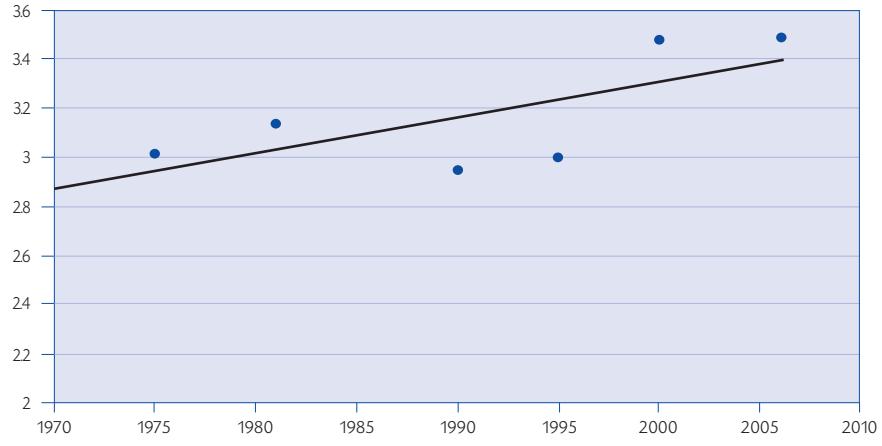
The happiness trend in India



Mean happiness in India 1975-2006 (1 = not at all happy, 4 = very happy)

Figure 14.B.

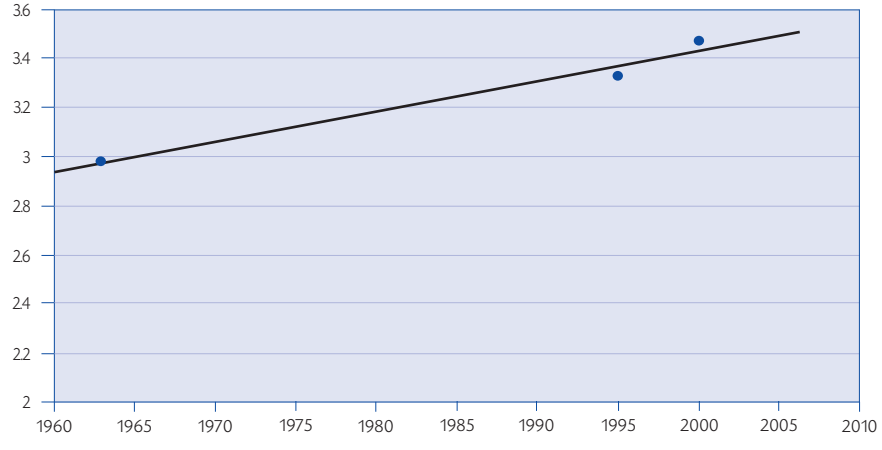
The happiness trend in Mexico



Mean happiness in Mexico 1975-2006 (1 = not at all happy, 4 = very happy)

Figure 14.C.

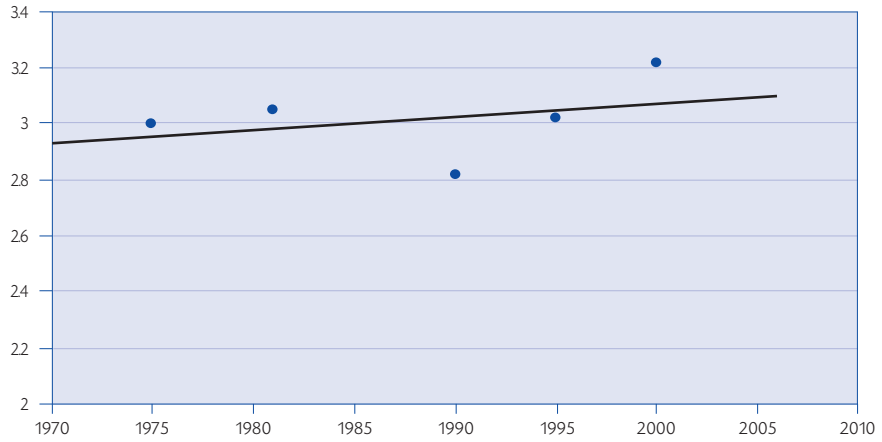
The happiness trend in Puerto Rico



Mean happiness in Puerto Rico 1975-2006 (1 = not at all happy, 4 = very happy)

Figure 14.D.

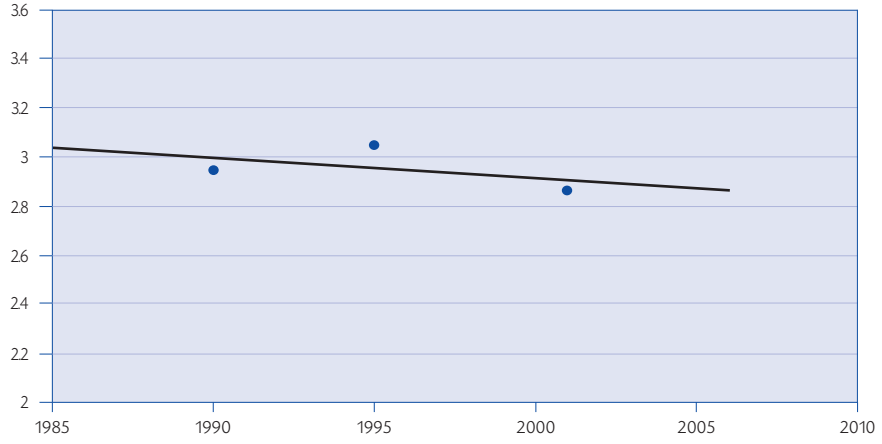
The happiness trend in South Africa



Mean happiness in South Africa 1981-2006 (1 = not at all happy, 4 = very happy)

Figure 14.E.

The happiness trend in China



Mean happiness in China 1990-2006 (1 = not at all happy, 4 = very happy)

A note on statistical power

The dispute over the long-run income-happiness gradient revolves around the magnitude of the correlation coefficient and its statistical significance. A number of authors have underlined that there is less statistical power in long-run series of well-being than in the cross-section, due to the smaller standard deviation. With less variation to explain, it is difficult to obtain statistically significant correlations.

Hagerty and Veenhoven (2000, p.5) for instance, note that "*the standard deviation in GDP* per capita *in the cross section from Diener and Oishi was about $8,000, whereas the standard deviation in Hagerty time-series (for the same countries) was only about 0.25 of that ($2,000)… within a country in 25 years.*" Hence, the statistical power to detect the effect is lower in time-series work. Equally, Kenny (2005), using data on 21 transition and developed countries, found a standard deviation in happiness over time within countries of 0.28 on average, as compared to a standard deviation of average scores across countries of 0.65 (p.212). Layard *et al.* (2010, p.161), using Eurobarometer time series for 20 Western European countries, also report an average standard deviation of national happiness scores over time of 0.2, as compared to an average of 0.5-0.6 in the individual cross-sections.

We calculated the standard deviation in happiness and life satisfaction in the World Values Survey cross-sections from 1981 to 2007. The average standard deviation within a **cross-section** (250 observations) is 0.67 for happiness (4-point scale) and 2.14 for life satisfaction (10-point scale). But the standard deviation of average national happiness **across countries** is 0.28 for happiness and 1.04 for life satisfaction. Finally, the standard deviation of national happiness **over time** fluctuates around 0.1 for happiness and from 0.13 to 0.41 for life satisfaction. In other words, the variability of subjective well-being measures is much lower in time-series than in the cross-sections within countries and across countries. The implication is that the difference between cross-sectional versus time-series correlation coefficients is difficult to interpret.

In summary, the long-run relationship between GDP growth and subjective well-being is still a subject of some controversy. As pointed by Stevenson and Wolfers (2008), one cannot reject the null that the correlation coefficient is equal to zero, but this does not mean that one can reject the null that it is greater than zero. The nature of the long-run relationship between GDP and well-being is far from being firmly established.

3.1.5. Subjective well-being and the business cycle

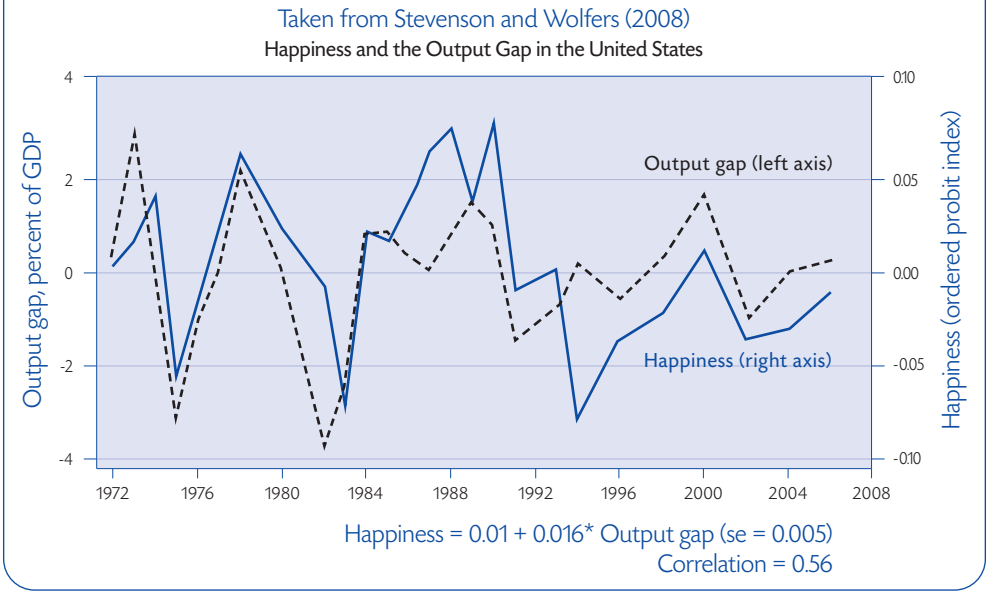
One of the reasons why it is difficult to admit no correlation between income and well-being is that this appears in sharp contradiction to the undisputed welfare effect of the business cycle.

There is first of all considerable consensus that recessions make people unhappy. Di Tella *et al.* (2003) showed that macroeconomic movements, in particular unemployment, inflation and the volatility of output, exert strong effects on the happiness of nations. The negative impact of volatility on subjective well-being was also established by Wolfers (2003). A powerful illustration of the business cycle-happiness correlation is given in Figure 15.A, taken from Stevenson and Wolfers (2008), which shows the spectacular parallel dynamics of the output gap and the average happiness in the United States from 1972 to 2008.

This does not mean that the influence of the business cycle can be equated with the influence of long-run growth, however. It is indeed easy to imagine happiness and the business cycle fluctuating around a flat long-run trend. While it is uncontroversial to say that happiness rises in booms and falls in busts, the key question is whether four percent growth in GDP *per annum* (for example) will produce a happier society in the long run than one percent GDP growth per annum.

Figure 15.A.

Happiness and the business cycle



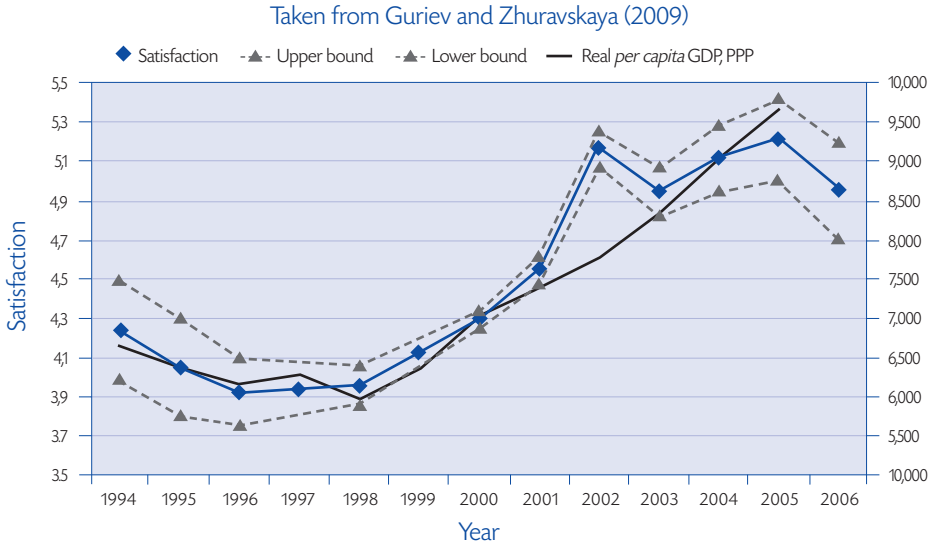
Source: General Social Survey, 1972-2006; Bureau of Economic Analysis.

Note: *"Output gap" is the difference between real GDP* per capita *and its trend, estimated using a Hodrick-Prescott filter on annual data on the logatitm of real GDP per capita* with *the smoothing parameter set to 6.25. Happiness data are aggragated into a happiness index by running an ordered probit regression of happiness on year fixed effects. See figure 17 for wording of the question. See text for details of the sample.*

One particular episode which is often considered as an illustration of the correlation between income fluctuations and well-being, rather than between long-term growth and well-being, is the transition process in Central and Eastern European countries from socialism to capitalism. All of the work here recognizes the statistically significant correlation between the dynamics of GDP and that of subjective well-being. Figures 15.B and 15.C, taken from Guriev and Zhuravskaya (2008) and Easterlin and Zimmerman (2009), illustrate the concomitant evolutions in income and happiness in a number of transition countries. Similar evidence can be found in Sanfey and Teksoz (2008).

Figure 15.B.

Happiness and transition in Russia



Left-hand scale: Life satisfaction for an average individual from the panel regressions with individual fixed effects and other usual controls (95% CI). Right-hand scale: Real GDP *per capita* in PPP-adjusted 2000 US dollars.

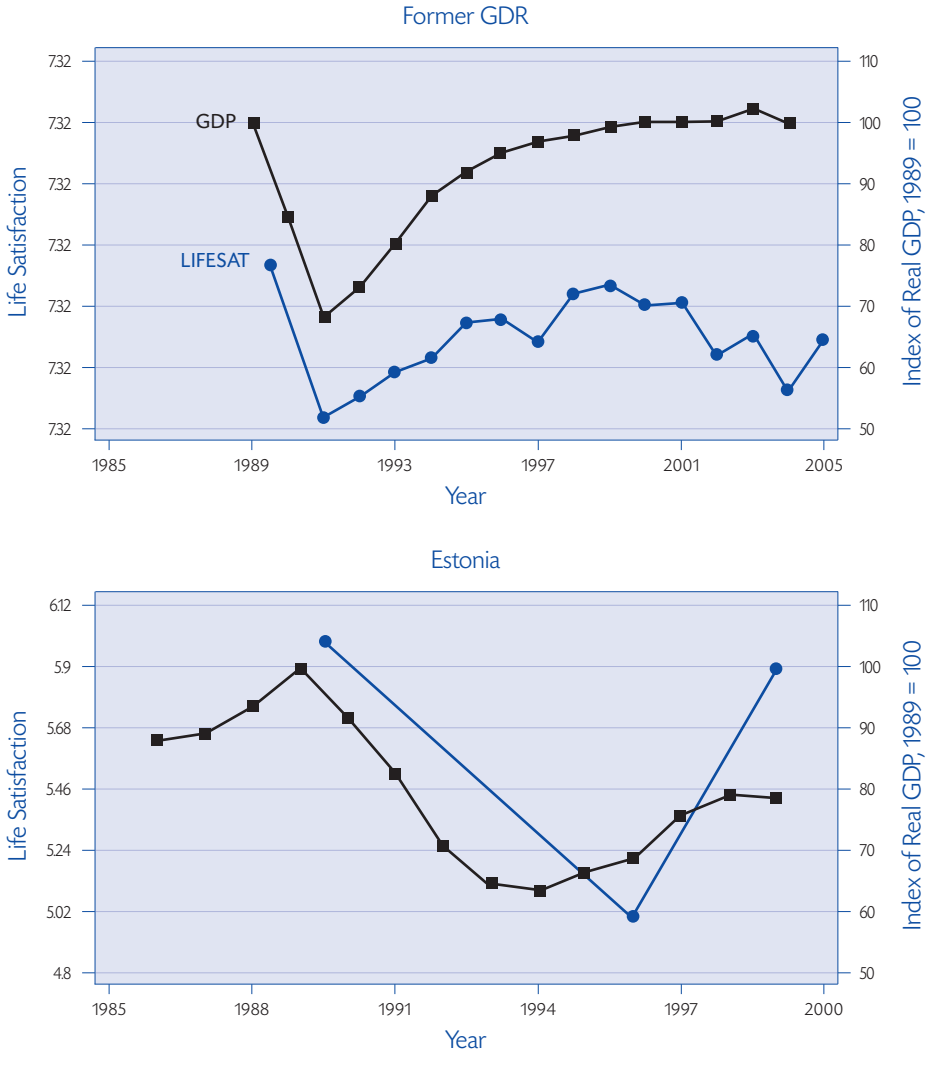
Sources: for satisfaction, the Russian Longitudinal Monitoring Survey; for GDP per capita*, the World Development Indicators database.*

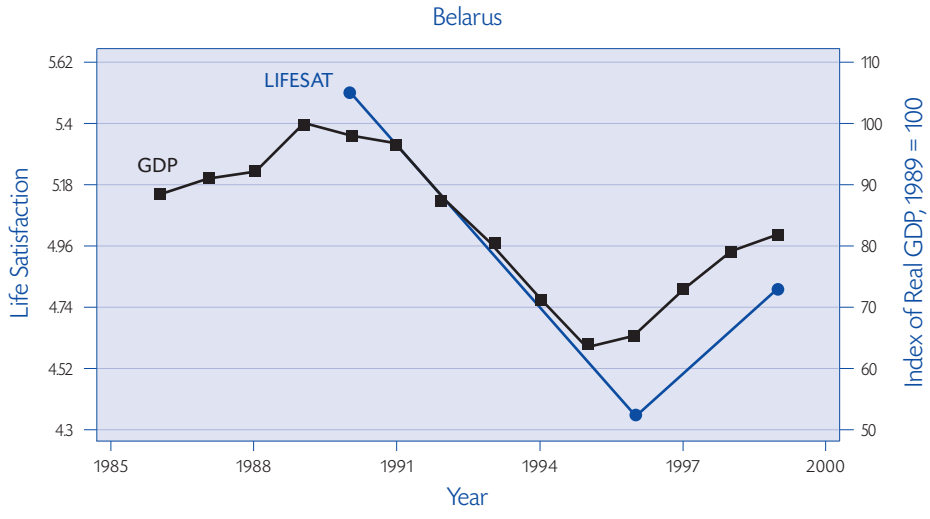
Figure 15.C.

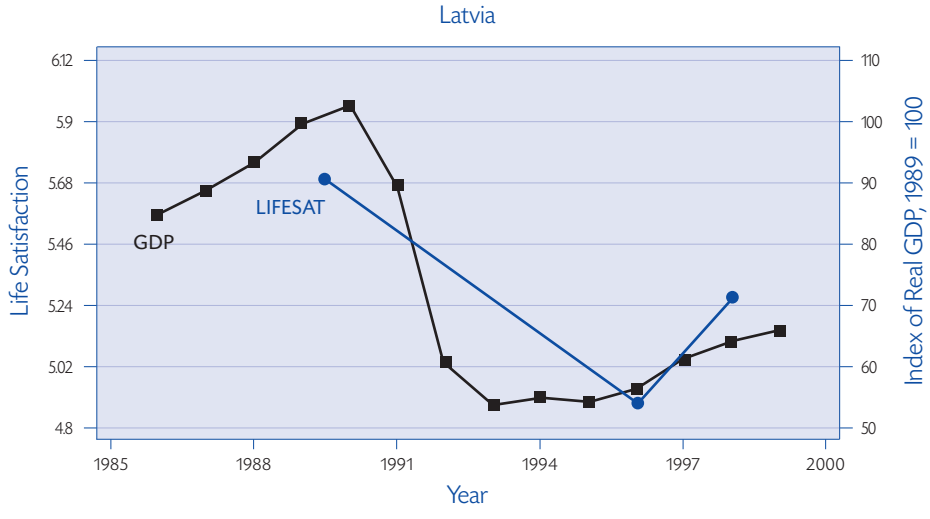
Happiness and transition in several countries

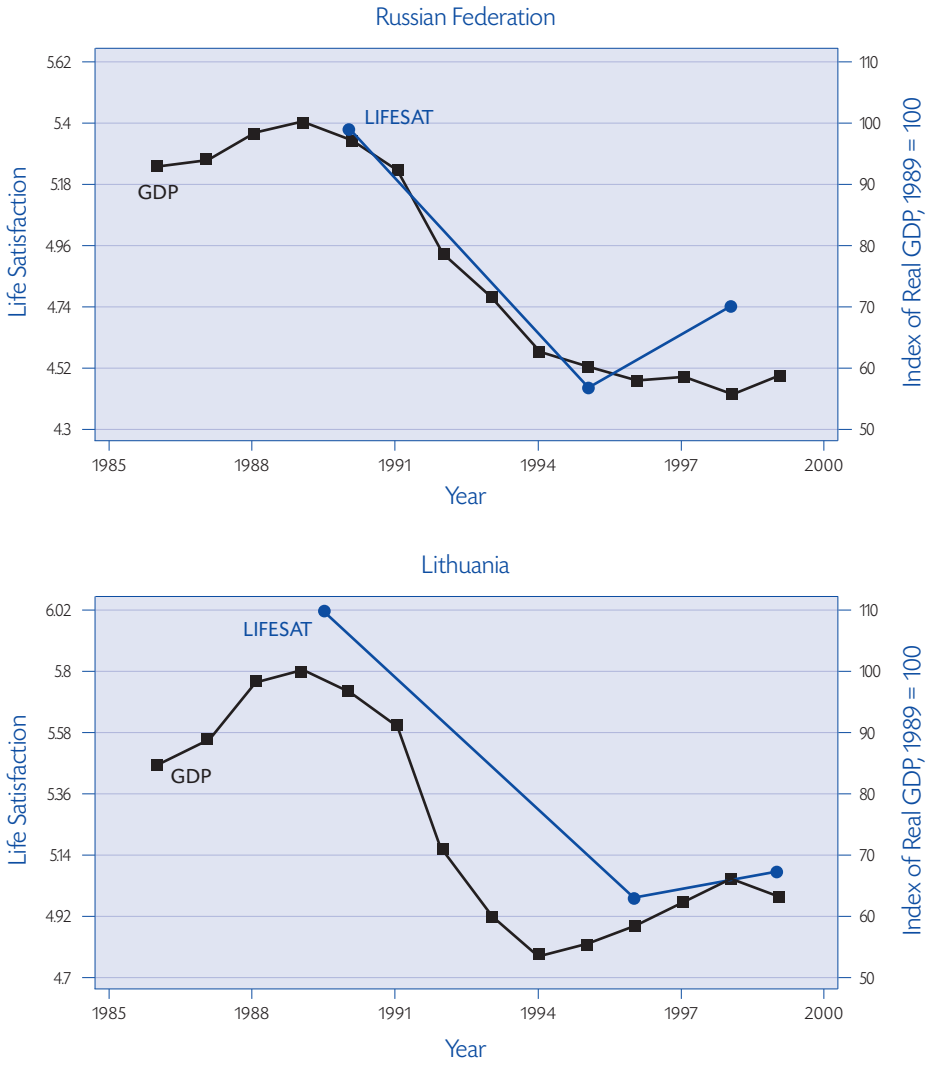
Taken from Easterlin and Zimmerman (2009)

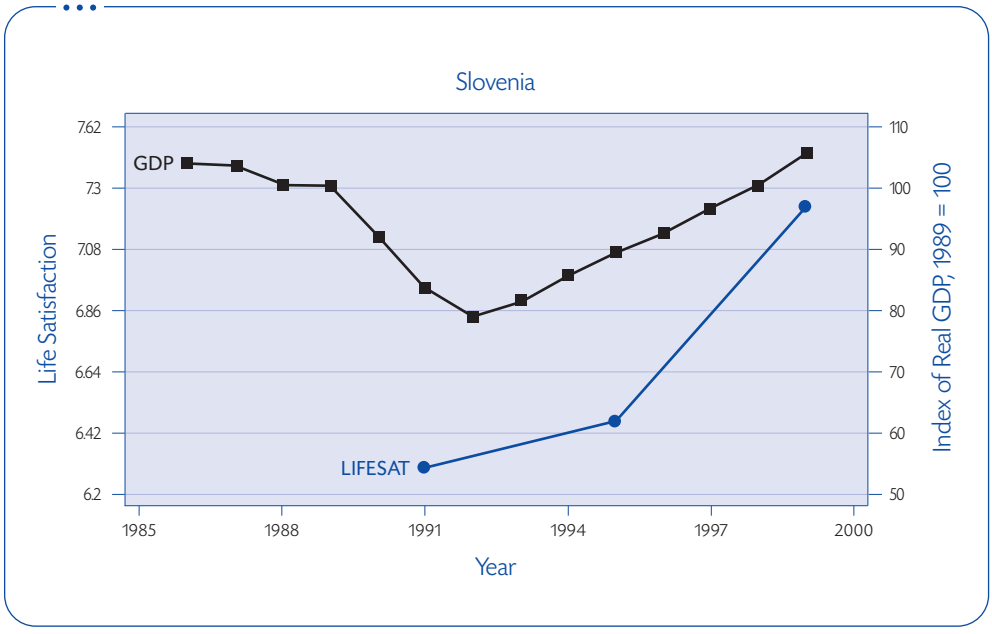
Life Satisfaction c. 1990, 1995, and 1999, and Index of Real GDP, Annually 1986-1999[a]




Belarus
LIFESAT
GDP
Life Satisfaction
Index of Real GDP, 1989 = 100
Year
5.62
5.4
5.18
4.96
4.74
4.52
4.3
110
100
90
80
70
60
50
1985
1988
1991
1994
1997
2000


Latvia
LIFESAT
GDP
Life Satisfaction
Index of Real GDP, 1989 = 100
Year
6.12
5.9
5.68
5.46
5.24
5.02
4.8
110
100
90
80
70
60
50
1985
1988
1991
1994
1997
2000


Russian Federation
Life Satisfaction
5.62
5.4
5.18
4.96
4.74
4.52
4.3
GDP
LIFESAT
1985
1988
1991
1994
1997
2000
Year
110
100
90
80
70
60
50
Index of Real GDP, 1989 = 100
Lithuania
Life Satisfaction
6.02
5.8
5.58
5.36
5.14
4.92
4.7
LIFESAT
GDP
1985
1988
1991
1994
1997
2000
Year
110
100
90
80
70
60
50
Index of Real GDP, 1989 = 100



Source: Real GDP, Economic Commission for Europe, 2003.
For Former GDP, GDP 2003 on is extrapolated from 2002 via real household income from GSOEP. Life satisfaction.

[a] *Former GDR, 1989-2005.*

However, these trends are qualified as short term by Easterlin and Angelescu (2009, p.14), who warn that one should avoid "*confusing a short-term positive happiness-income association, due to fluctuations in macroeconomic conditions, with the long-term relationship. We suggest, speculatively, that this disparity between the short and long-term association is due to the social psychological phenomenon of 'loss aversion'*".

However valuable the interpretation in terms of loss-aversion, it is perhaps surprising that transition is considered to be only a short-term phenomenon. In a way, transition is the best example of regime change that we can think of. It is a deep and irreversible structural transformation, not a short-lived phenomenon. It shares the essential features of development, including the take-off period and the profound qualitative and institutional changes. Hence, whether transition should be treated as a short-term or a long-term phenomenon remains an open question. Only the passage of time will enable us to see whether the increase in subjective well-being continues with GDP growth, stagnates at a certain point, or falls back down to the initial (1990) level.

3.2. Explanations related to growth itself: channels and negative side-effects

The flatness of happiness curves is therefore consistent with GDP growth not yielding higher well-being over time. More generally, it may suggest that whatever changes a country experiences over time have no long-run effect on individual average happiness. If this is true, the prospect is dark for developing countries, which are locked in at their current low level of happiness. The message is also very discouraging for public policy in general: if happiness cannot be raised in the long run, not only should growth be abandoned as an objective, but so should any other public policy measure.

Before jumping to these radical conclusions, the two next sections discuss possible explanations of the flatness of the happiness curve. A first series of explanations pertain to the nature of growth itself, *i.e.* the channels of growth and the fact that growth is accompanied by negative externalities (pollution, inequality, etc.) that cancel out its subjective benefits. The second series of explanations cover social and psychological processes, such as comparisons and adaptation, that may well reduce the happiness benefits of growth.

3.2.1. Quality of Life: channels from GDP growth to happiness

Statistical estimates of subjective well-being most often include time and/or country fixed effects, as well as other controls that are introduced in order to pick up any changes in the demographic composition of the population (in terms of age, occupation, health, number of children, etc.). Some analyses also control for political variables such as democracy, gender equality, trust, etc. However, in terms of the empirical strategy retained for the estimation of the relationship, there is always a trade-off between controlling for variables that reflect the channels *via* which the phenomenon under consideration works, and not controlling for omitted variables and obtaining a biased measure of the relationship. For example, in the context of the current question of growth and well-being, a well-being regression that controlled for both GDP and the positive side-effects (or channels) of growth runs the risk of concluding that growth does not matter for well-being. Indeed, we expect growth to bring about higher well-being not only *via* greater purchasing power (income), *i.e.* through higher consumption, but also *via* other transformations (education, health, etc.) which accompany the growth process. Controlling for these latter transformations may render GDP itself

insignificant in a well-being equation, but that does not mean that greater income does not produce greater happiness, it rather means that we have identified the different processes *via* which income produces well-being.

Greater income *per capita* always comes with increased productivity, which means a greater choice in time-use for those who are concerned. As argued by Sen (2001), it is because it enhances the freedom of choice (by enlarging their set of capacities) that growth is expected to raise people's well-being. Identically, GDP growth is known for being associated with demographic transitions in developing countries. This is certainly "*a revolutionary enlargement of freedom for women*", as put by Titmuss (1966, quoted by Easterlin and Angelescu 2007, p.9), and a rise in the education and resources for self-development that children can count on. Growth also comes with higher life expectancy, reduced child mortality and child underweight (see for instance Becker, Philipson and Soares, 2005; Easterlin and Angelescu, 2007). Finally, it is well-known that democracy and development go hand in hand, even if the direction of causality is not as clear as was believed in the 18th century (*e.g.* by Montesquieu, Steuart and Hume). Lipset (1959, p.80), for example, claims that: " *industrialization, urbanization, high educational standards and a steady increase in the overall wealth of society [are] basic conditions sustaining democracy.*" Without inferring any causality, we can observe the statistical association between GDP growth and progress in terms of political freedom and human rights. With respect to the empirical strategy, any attempt to capture the global effect of GDP growth on subjective well-being should not control for any such variables which represent the channels of transmission. It is likely regrettable that much of the work on the GDP growth-happiness relationship does indeed include such controls.

The following sections review the available evidence on the correlation between GDP growth and such quality of life indicators. These latter are measures of the non-income quantitative and qualitative dimensions that constitute the channels from income growth to well-being.

Cross-section correlation between GDP growth and quality of life indicators

Easterlin and Angelescu (2007) illustrate the sizeable positive correlation in cross-section data between a number of quality of life indicators and GDP *per capita* across countries at different levels of development. The clear upward slopes relate subjective well-being to quantifiable factors, measured on continuous scales. These latter include food, shelter, clothing and footwear, energy intake, protein intake, fruit and vegetables, radios, cars, TV sets, mobile phone subscriptions, Internet users, urban population, life expectancy at birth, gross education enrolment rate, and the total fertility rate. These kinds of relationships have been documented by a considerable number of other authors, including Inglehart and Welzel (2005); Inglehart *et al.* (2008); Layard *et al.*, 2010; Di Tella and MacCulloch (2008); Becker *et al.* (2005).

Along analogous lines, some authors have insisted on the relationship between subjective well-being, on the one hand, and procedures, governance and institutions, democratic and human rights, tolerance of out-groups, gender equality, on the other (for example, Barro, 1997; Frey and Stutzer, 2000; Inglehart and Welzel, 2005; Schyns, 1998; Inglehart *et al.*, 2008).

Time-series correlation between GDP growth and quality of life indicators

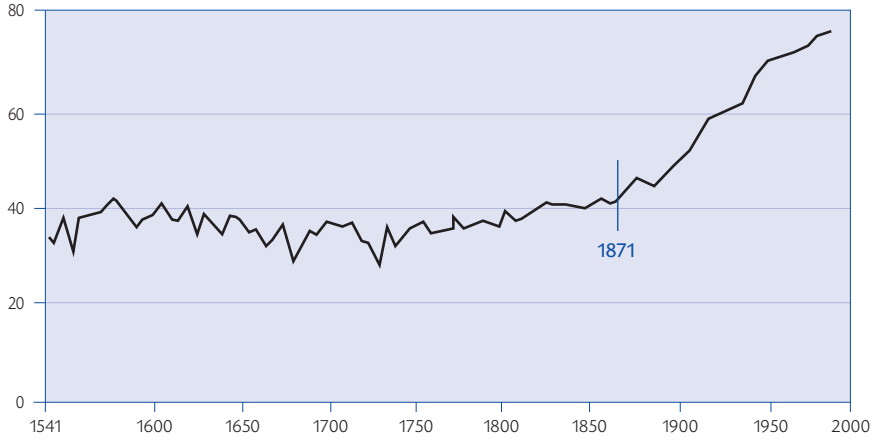
Figure 16 illustrates the spectacular take-off of life expectancy in England and Wales in the 19th century. More generally, Easterlin and Angelescu (2007) provide a detailed account of the progress in the different dimensions of quality of life over time, in a large set of developed and emerging countries. They document the different dimensions of changes in the quality of life during "*modern economic growth*". The latter is defined as a "*rapid and sustained rise in real output per head and attendant shifts in production technology, factor input requirements, and the resource allocation of a nation*", where "*rapid and sustained*" is defined as being equal to at least 1.5 percent per year (Easterlin and Angelescu, 2007, p.2).

Figure 16

The take-off in life expectancy

Taken from Easterlin and Angelescu (2007)

Figure II B-3. Life expectancy in England and Wales since the sixteenth century



Source: 1541-1871, Wrigley and Schofield (1981, p. 230); 1871 to 1945-47, Keyfitz and Flieger (1968, pp 36-9); 1950-55 to 1990-95, United Nations (1995).

Easterlin and Angelescu document the turning points in GDP growth and other indicators of the quality of life. Although both variables move in parallel, they insist that the dates of their respective take-offs do not systematically coincide. Qualitative indicators sometimes lag behind and sometimes lead the date of GDP take-off. "*If social and political indicators of QoL are, at present, positively associated with GDP per capita, it is often because the countries that first implemented the new production technology underlying modern economic growth were also the first to introduce, often* via *public policy, new advances in knowledge in the social and political realms*" (Easterlin and Angelescu, 2007,

p.21). Whether the co-movements between growth and quality of life indicators represent a causal relationship is controversial and difficult to establish (see also Easterly, 1999). However, it is undeniable that overall there is no progress in quality of life without GDP growth.

In their provocative paper, *Is Growth Obsolete?*, William Nordhaus and James Tobin (1973) advocated for an alternative indicator, integrating leisure, household work and costs of urbanization, and constructed a "Measure of Economic Welfare". However, this index turned out to grow in a way that was similar to GDP over the period under study, albeit more slowly. This, to our knowledge is a universal observation. Pritchett and Summers (1996), for example, note that "wealthier is healthier" in the long run. Using time-series data from a variety of countries, they find that: "*The long-run income elasticity of infant and child mortality in developing countries lies between 0.2 and 0.4.*" This implies that "*over a half a million child deaths in the developing world in 1990 alone can be attributed to the poor economic performance in the 1980s.*"

In summary, GDP growth goes hand in hand with a series of quantitative and qualitative non-monetary improvements in quality of life. These constitute the channels from growth to well-being that we argue should not be controlled for in the statistical analysis of the former relationship.

3.2.2. Negative side-effects of growth

The flatness of the GDP-happiness graphs may be due to the negative influence of some side-effects of growth, such as pollution, income inequality, work stress, and so on. The influence of these "omitted variables" could then well hide the positive influence of GDP growth on subjective well-being in econometric analyses (see Di Tella and MacCulloch, 2008).

The most widely discussed negative side-effects of growth are: inequality, crime, corruption, extended working hours, unemployment, pollution and other environmental degradation (as measured by SOx emissions, for example). These are discussed in Di Tella and MacCulloch (2003, 2008). Kenny (2005) also emphasises the social cost of economic transformation, and the ensuing shift from local to global relative income concerns. The impact of urban concentration and sub-urbanization is not so clear-cut, however. Easterlin and Angelescu (2007) also underline the effects of carbon dioxide emissions and fat intake (obesity and blood pressure). Clark and Fischer (2009) provide a useful summary of the macro-economic correlates of life satisfaction in OECD countries.

Among the list of usual suspects, income inequality occupies a particular place. In the first place, the relationship between income inequality and subjective well-being has been the subject of a considerable body of work, much of which has concluded to a negative correlation (see Senik, 2009, for a survey; and Clark *et al.*, 2008, and Alesina and la Ferrara, 2008, for surveys of the self-reported demand for income redistribution). Income inequality will reduce well-being if people dislike it as

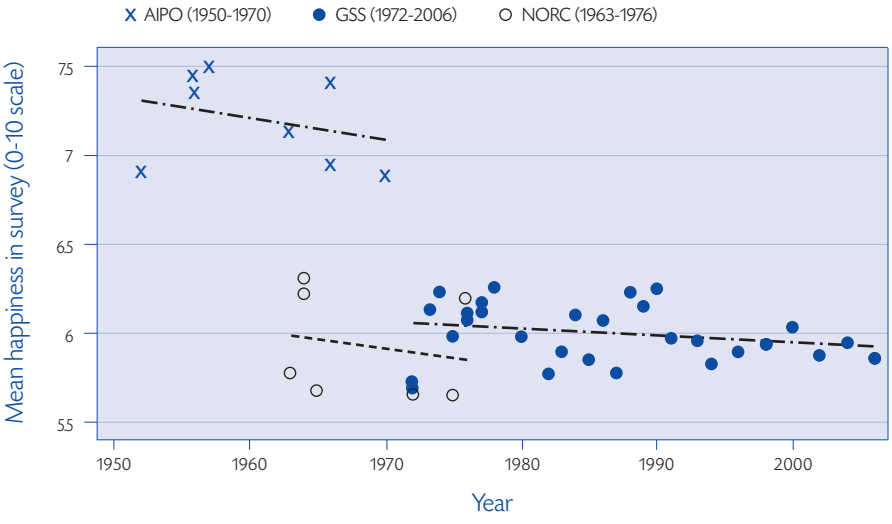
such (although, on the other hand, it will be associated with higher well-being if it is interpreted as reflecting a greater scope of opportunities: see Alesina *et al.*, 2004). However, it can also exert a mechanically negative effect on average SWB, due to the concave relationship between income and SWB (see Stevenson and Wolfers, 2008). However, this mechanical effect does not seem to be sufficient to explain the flatness of the curve. As illustrated by the different panels of Figure 17 (taken from Layard *et al.*, 2010, p.142), income inequality increased sharply from 1970 to the end of the 2000s, but average happiness has remained flat. In addition, the income of the upper quintile of the income distribution has risen, but the happiness scores within this quintile have not. Hence, even for highest income quintile, the happiness curve has remained flat in the USA.

Figure 17

GDP growth, inequality and happiness

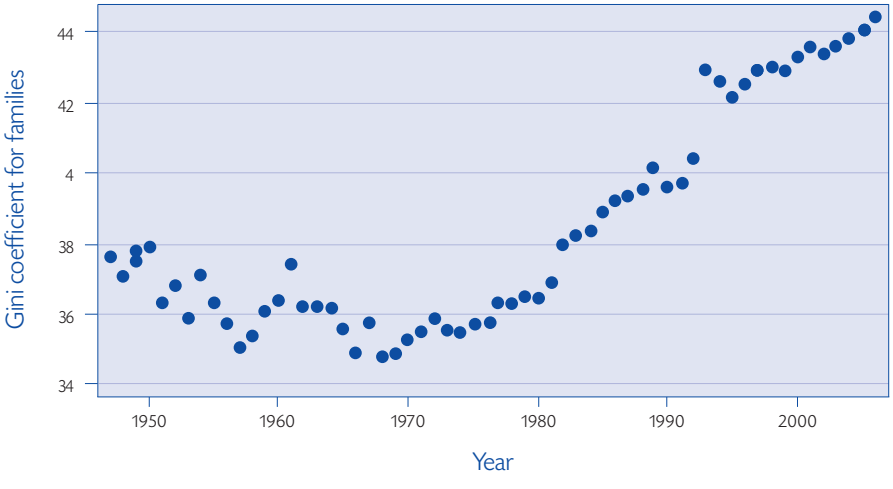
Taken from Layard, Mayraz and Nickell (2010, p.142)

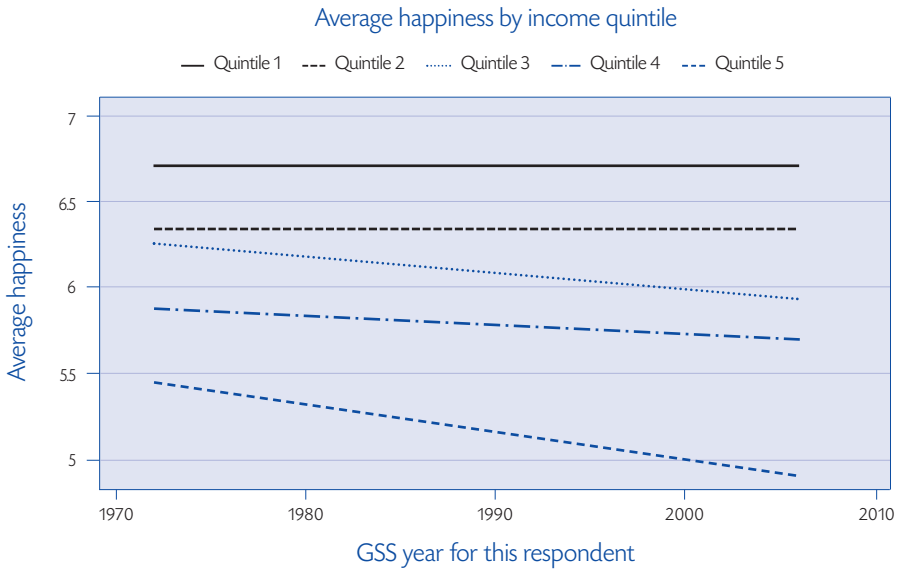
Reported happiness over time in the United States

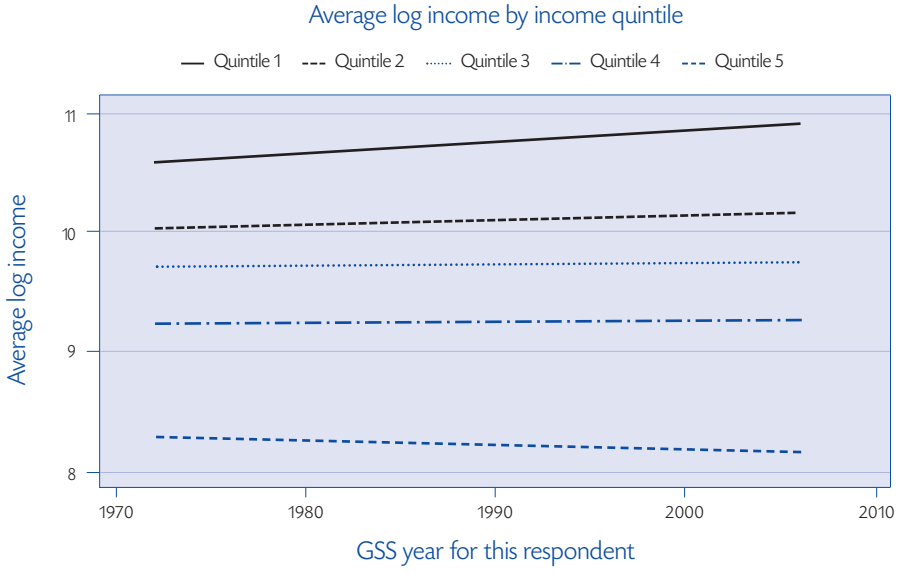


AIPO is the American institute of public Opinion (later Gallup).
NORC is the National Rechearch Center, and administrator of GSS.
GSS is the General Social Survey.
Original questions asked on a 3-point scale.

Family income inequality over time







Source: GSS 1972-2006. Income is adjusted for the number of adults in the household.

One important note that can be made here is that many of the negative externalities of growth seem to exhibit an inverted U-shape, *i.e.* they increase in the initial stages of development and then subsequently fall in the later stages. Income inequality, pollution, long hours of work, poor working conditions, etc. are phenomena that initially seem to have grown in importance with income growth, but which have then been attenuated at some point in high-income countries. This is not only the result of purely mechanical forces, but also of public policy: this is an important point to make in the context of developing countries.

Should these negative factors then be taken into account when evaluating the effect of GDP growth on well-being? This an open question. If these negative side effects constitute inevitable companions to growth, then the answer is Yes: they have to be counted negatively in the welfare accounting of growth. However, if these side-effects can potentially be attenuated or suppressed by public policy, then they are not necessarily intimately linked with higher income, and as such their well-being effect can be removed from the welfare effect of growth.

3.3. Explanations related to the happiness function itself (human beings are social animals)

3.3.1. Income comparisons

One simple explanation of the lack of any long-run relationship between income and well-being is that this does not reflect that there is something wrong with growth *per se*, but rather that this reflects the very structure of individual well-being functions. The broad idea is that income does not bring well-being in a vacuum, but is rather intensely social, in that it is evaluated relative to some benchmark, reference or comparison level of income. There are many synonyms for the latter: this can be thought of as what is normal in the society, or what is fair. Forgetting about the other determinants, we can then write the relationship between utility and income as:

(1) $U_{it} = U(y_{it}, y_{it}^*)$

The well-being of individual *i* at time *t* rises with their own income, y_{it}, but falls with the level of comparison income, y_{it}^*. Comparison income acts as a deflator with respect to own income here, in the sense that the higher it is the less good the individual's own income looks. Much of the empirical literature exploring this relationship has explicitly parameterized the well-being function as a function of both y_{it}, and y_{it}/y_{it}^*. If the income effect of income on well-being is mostly absolute, so that in the absence of the above-mentioned externalities greater GDP will increase individual well-being, then the second term will play only a minor role. On the other hand, if income comparisons are very important, so that most of the effect of income works through how well I am doing compared to some reference group, then it is the second term that will be preponderant. If it is mostly relative income (y_{it}/y_{it}^*, which is homogeneous of degree zero) that matters, then, answering Dick Easterlin's 1995 question, "raising the incomes of all" will not "increase the happiness of all".

Distinguishing between these two scenarios has been the goal of a considerable amount of empirical work over the past fifteen or so years. A variety of different empirical approaches across various disciplines have been mobilized to answer the question of how much income comparisons matter in the determination of well-being. All of this work has had to set out *a priori* exactly to whom or to what individuals are thought to compare themselves: this has included the individual's spouse, to people with the same characteristics as the individual, those in the same region, other participants in experiments, hypothetical individuals, or even a measure of the individual's expected income. Some of the key findings in developed countries are described below.

Evidence in developed countries

One direct approach to the question of income comparisons has been to estimate

well-being regressions in which both the individual's own income and the comparison income level appear: these are the empirical counterpart to equation (1) above. This literature has appealed to different datasets (in terms of countries and years), different measures of well-being (job and life satisfaction being the most predominant), and various measures of comparison income, y_{it}^*. The typical finding is that own income is positively correlated with well-being, but that the correlation with others' income is negative.

Clark and Oswald (1996) use the BHPS to calculate the income of "people like me" from a wage equation, and show that this is negatively correlated with individual job satisfaction. Own income attracts a positive coefficient, and the sum of the two estimated income coefficients is zero: pay rises for everyone have no effect on satisfaction. More recent work along the same lines using, respectively, German and American data is Ferrer-i-Carbonell (2005) and McBride (2001). Vendrik and Woltjer (2006) extend the analysis of the German GSOEP data in this respect, by considering asymmetric reactions to gains and losses (relative to the reference group).

An alternative measure of y_{it}^* is at the local level: what do my neighbours earn? Both Blanchflower and Oswald (2004) and Luttmer (2005) calculate regional average income in US data, and show that this is negatively correlated with respondents' well-being: an individual earning $40,000 per year is happier in a poorer than a richer region. However, at the very local level of a few hundred metres, Clark *et al.* (2009) find that in Danish panel data, conditional on my own income and local median income, my satisfaction is strongly positively correlated with my rank in the local income distribution. Other work here has considered comparisons to the income of the individual's work colleagues (Brown *et al.*, 2006), partner (Clark, 1996) and parents (McBride, 2001).

Running well-being regressions is only one way of addressing the question of income comparisons. One early method (the first published contribution being Van Praag, 1971) is that of the Welfare Function of Income. Here individuals assign income levels (per period) to verbal labels (such as excellent, good, sufficient and bad): these stated values form the basis of individual-level regressions estimating a lognormal Welfare Function of Income. The resulting individual estimated means (μ) reveal which individuals require greater income in order to be satisfied. Comparison income is introduced into the analysis, typically as average income over age, education and other characteristics. The regression results (for example, Van de Stadt *et al.*, 1985) show that, given own income, the higher is reference group income, the more money individuals say they need to reach a given verbal well-being level, which is consistent with income comparisons.

Separate evidence on comparisons is found in experimental economics. In Zizzo and Oswald (2001), experimental participants paid out of their own winnings in order to burn the money earned by other participants. An alternative approach is to ask individuals to choose between hypothetical outcomes, as in Alpizar *et al.* (2005), Johannsson-Stenman *et al.* (2002) and Solnick and Hemenway (1998). A typical income choice is as follows:

A: Your current yearly income is $50,000; others earn $25,000.

B: Your current yearly income is $100,000; others earn $200,000.

The key here is that one choice has a greater absolute return while the other is more advantageous in relative terms. In line with the experimental work, there are strong positional concerns over income, in that individuals choose *A* over *B*. While the above example is couched in terms of income, the same method can be used to compare the degree of comparisons across domains. For example, relative concerns in Alpizar *et al.* are stronger for cars and housing, and weaker for vacations and insurance.

A recent randomized experiment was set up by Card *et al.* (2010), showing evidence of relative concerns among employees of the University of California when they had access to Internet information about the wage of their colleagues.

Last, we can appeal to recent neurological work. Fließbach *et al.* (2007) use MRI techniques to measure the brain activity of pairs of individuals engaged in identical guessing-game tasks. Each individual's monetary reward for a correct guess was announced to both subjects, and these rewards were varied. In some conditions a correct guess by a participant earned 60 points; in other conditions the subject's guess earned 60 and the other's correct guess earned 30, or 60 and 120. As such, the individual's relative payoff for a correct guess changed, while keeping the absolute reward fixed. Blood oxygenation analysis showed that brain activity in the ventral striatum was increased with relative income. Related work in this area appears in Takahashi *et al.* (2009).

Evidence in LDCs

The majority of the work on income comparisons and individual well-being has covered OECD countries. However, the increasing availability of data, including subjective questions undoubtedly allied with the increasing interest that researchers have in these issues, have produced a small but growing number of pieces of evidence regarding the correlates of individual well-being in poorer countries. The key question that we want to answer here is whether positional concerns are less important in poorer countries: Are comparisons luxuries?

Regarding the direct estimation of individual well-being, Graham and Felton (2006) have replicated the finding of a negative effect of regional income on individual well-being across 18 Latin American countries. Kuegler (2009) analyses self-collected data on 400 Venezuelans in 2005, and shows that those who say that they are better-off than their own siblings report higher life satisfaction. This is consistent with relative income effects in a relatively poor country. The strength of this correlation depends on the individual's own characteristics, being stronger for respondents with above-median incomes and those who work in higher-rank professions. Stark and Taylor (1991) present indirect evidence of the role of income comparisons by looking at the decision to migrate. Using Mexican data, they show that relative deprivation is a significant predictor of Mexico-US migration.

Castilla (2010) also considers Mexican data, including information on subjective poverty (whether the respondent's income is sufficient for their needs) and income satisfaction. Relative concerns are introduced by considering these two welfare measures as a function of both own expenditure and the respondent's evaluation of their own income relative to people with whom they live, to how much they aspired to have at this stage of their lives,

and relative to the income they earned three years ago (all three of which are measured on a seven-point scale). The empirical results show that welfare rises with own expenditure, but falls with income relative to others and income relative to aspirations. The results with respect to past income are significant only in the life satisfaction equation and when the individual reports being worse off than three years ago (which is consistent with loss-aversion).

Rojas and Jiménez (2007) also appeal to Mexican data to show respondents' subjective poverty evaluations are partly determined by the gaps between own income, on the one hand, and comparison and aspired income levels, on the other. Comparison income is measured directly by asking about the income gap "with respect to those you usually compare yourself to". Guillen-Royo (2010) analyses small sample data from seven communities in Peru, and shows that satisfaction with a number of different life domains is positively correlated with own expenditure but negatively correlated with average community expenditure. Last, Rojas (2010) uses data from 20 Latin American countries found in the 2007 Gallup survey. Two measures of individual well-being, the ladder question of worst to best possible life and satisfaction with standard of living, are related to both own income and the average income in the reference group (defined by age, sex and country). The empirical results show that well-being rises with the log of own income but falls with the log of comparison income. In the case of satisfaction with standard of living, the coefficients on the two variables are equal and opposite, suggesting that a rise in everyone's income would leave no-one in Latin America better-off.

Moving from Latin America to Asia, there has been a spate of recent work on the determinants of well-being in China, some of which has appealed to the notion of reference income. Appleton and Song (2008) conclude that the life satisfaction reported by urban Chinese is affected by status considerations, and Smyth and Qian (2008) analyse data from 31 Chinese cities in September 2002, finding that the log of average monthly income in the city in which the respondent lives is negatively correlated with happiness, controlling for own income. Gao and Smyth (2010) appeal to two different datasets to present some evidence that job satisfaction is negatively related to reference group income, where this latter is either average income in the firm in which the respondent works, or the predicted income of "people like me" (as in Clark and Oswald, 1996).

Recent work by Cojocaru (2010) appeals to cross-section 2007 data from the Living Standard Measurement Survey (LSMS) in Tajikistan. He finds a mostly insignificant effect of regional income on individual life satisfaction, but suggests that this might reflect the fact that the wrong reference group is being used. When, however, a qualitative variable is used which measures the individual's evaluation of their household's welfare relative to that of their neighbours, strong effects are found in the expected sense: those who rank their household relatively lowly compared to their neighbours report lower levels of life satisfaction, controlling for the household's own expenditure.

Fafchamps and Shilpi (2008) consider a direct measure of relative utility in a developing country by analysing the answers to a question on consumption adequacy in Nepalese data. Consumption adequacy rises with own

income (but falls with the distance to the nearest market). Critically, conditional on these and other control variables, consumption adequacy also falls with reference group consumption, as in a relative utility model. Here reference group consumption is defined in a geographical way as the mean or median consumption of other households living in the same ward as the respondent.

Carlsson *et al.* (2009) look at hypothetical preferences over different absolute and relative income situations (as used by Alpizar *et al.*, 2005) in India. They find that around half of the effect of income on well-being comes from some kind of status or relative income concern. Crucially, they note that this figure is around the same as that found in rich countries. They moreover note that low caste and low income respondents seem to be more sensitive to relative income.

John Knight has authored a series of papers using Chinese data from the national 2002 Chinese Household Income Project (CHIP) survey. Unusually, this survey included not only questions on subjective well-being but also asked direct questions about who individuals considered as their reference group. Knight *et al.* (2009) appeal to cross-sectional information on 9,200 households in China.The authors first show that comparisons in China are local, in that 70 percent of individuals see their village as their reference group. Further, conditional on both own and village income, those who report that their own income was much above the village have higher happiness scores. Knight and Gunatilaka (2010*a* and 2010*b*) also emphasize the importance of relative income rather than absolute income, and the role of changing reference groups, in Chinese data. Mishra *et al.* (2010) show that reporting an income below that of a self-reported reference group is associated with lower well-being for the Korean minority in China.

Well-being work using Chinese data has thus uncovered a number of pieces of evidence consistent with the presence of income comparisons in a developing country. This is consistent with the results in Brown *et al.* (2010), who do not measure well-being directly, but instead appeal to the literature that has analysed conspicuous consumption in developing countries. They use data from a Chinese household panel, and show that spending on funerals and gifts is consistent with status-seeking behaviour. Last, Fließbach and co-authors followed up their 2007 work by running the same relative income Neuro experiments in China (although the results have not yet been written up).

Turning to Africa, Kingdon and Knight (2007) consider the role of relative income in South Africa. The authors find evidence of negative relative income effects within race groups (whereby life satisfaction is lower the more others earn), but positive relative income effects within neighborhoods.[59]

Bookwalter and Dalenberg (2010) analyse South African SALDRU data from the early 1990s. They find no significant effect of local (cluster-level) income for Whites, but a positive and significant effect of others' income for non-Whites. However, similar to Cojocaru

[59] So that higher neighbourhood income is associated with greater satisfaction. This mirrors the finding in Danish small neighbourhood data in Clark *et al.* (2009).

(2010), dummy variables for one's own income compared to that of one's parents attract significant estimated coefficients consistent with income comparisons (with feeling less well-off than one's parents having a far larger absolute effect on satisfaction than feeling better-off than one's parents).

Ravallion and Lokshin (2010) appeal to large-scale 2004 household data from Malawi, which include measures of satisfaction with life and consumption expenditure. More unusually, the data also include measures of own subjective economic welfare, from respondents' answers to the question "*Imagine six steps, where on the bottom, the first step, stand the poorest people, and on the highest step, the sixth, stand the rich (show a picture of the steps). On which step are you today?*", as well as their assessment of the economic welfare of their neighbours and their friends. Ravallion and Lokshin model individual life satisfaction as a function of both own and local neighbourhood consumption, and as a function of both own and others' economic welfare. Although they argue that the results show that comparisons are not important for the majority of Malawians, others' consumption reduces individual life satisfaction in the urban sample, and there is some evidence of a negative effect of friends' economic welfare on those who report a relatively high level of own economic welfare.

On a smaller scale, Kenny (2005) uses data from a survey of 566 Tanzanian households, in which respondents report the amount of income necessary to be wealthy. Similar to the European results in Van Praag's work, it is shown that the average income in the area is one key determinant of what people consider to be a healthy income.

Akay and Martinsson (2008) use a cell-mean approach similar to that in Ferrer-i-Carbonell (2005) applied to 2004-2005 household survey data in Northern Ethiopia. They find no significant effect of reference group income on life satisfaction. A companion paper (Akay *et al.*, 2009) again looks at Ethiopia, but this time considers hypothetical preferences over absolute and relative income scenarios. The results here are that the choices of most Ethiopian subsistence farmers are based on absolute income alone. However, there are still an arguably considerable number of some of the poorest people in the world who take status considerations into account. Corazzini *et al.* (2010) use the same approach to compare the degree of relative income concerns across eight different countries. While they argue that there is a broad pattern of individuals in richer countries being more sensitive to relative income, it is striking that one of the most comparison-conscious countries in this respect is Kenya.

Absolute *versus* relative poverty

One of the reasons why we are interested in income comparisons, especially in the context of less well-off countries, is that they impinge on the concept of poverty. The distinction between poverty as an absolute lack and a relative lack goes back at least to Adam Smith: in the mid-18th century the Scots were not seriously deprived if they did not have shoes, whereas in England, only the truly destitute had no shoes. The stigma from being shoeless was therefore greater in England than in Scotland, because of the social norm that was attached to it. As such, the impact of a given lack on individual well-being may depend on the degree to which this lack is stigmatised in society, which itself is likely related to the incidence of the lack under consideration.

Moving back to income and appealing to equation (1) above, the critical distinction is then whether poverty is defined by an individual's income falling below a certain critical level, or whether other people's outcomes play a role. Absolute measures of poverty include the cost of minimum calorie intake line, the minimum consumption basket defining the poverty line in the US, and the World Bank's $1-a-day poverty line. Relative measures of poverty take context into account, such as the commonly used relative poverty line set at 60 percent of median income. The evidence of relative income concerns in low-income countries seems to constitute an argument in favor of measures of relative poverty.

Another important question that we are unable to answer to date, is whether relative concerns are less important, *i.e.* have smaller welfare effect in low income countries than in high-income countries. Income interactions can be thought about as some kind of luxury good, that come into attention only once survival is taken for granted. We have reviewed the evidence that relative concerns do exist in developing countries. But whether their importance is smaller than in developed countries remains an open question that would need specific data – maybe experimental data – to be answered. Analysing the data from the third wave of the European Social Survey, Clark and Senik (2010) focused on the answers to the question "*How important is it to you to compare your income with other people's incomes?*" across European countries. They found that this importance is greater in poorer countries than in richer countries, and that, within countries, this comparison is more often said to be important by poorer people. Comparisons are most often upward directed and people suffer more from upward-directed comparisons. This is consistent with the literature's general findings (see for example Ferrer-i-Carbonnell 2004, or Card *et al.* 2010). If this finding could be extended to poor countries, this would rule out the idea that income comparisons are a rich country phenomenon.

Knowing that local income comparisons matter for low income countries' citizens, one should consider the possibility that global income concerns may also be important, especially in view of the development of information and communication technologies. If the latter allow the inhabitants of low income countries to be aware of the life-style and consumption possibilities of high income country citizens, this is likely to generate feelings of relative deprivation. This might explain the steeper curve of the relation between GDP *per capita* and subjective well-being in developing countries (see introductory section). We are not aware of any direct evidence of global income concerns. One exception is Clark and Senik (2010), who noted that in the above-cited recent survey of Europeans, respondents who did not have Internet access were less subject to income comparisons.

The most radical view about the importance of income comparisons would lead to the conclusion that it is only because they compare to others that the richer inhabitants of the globe are more happy and the poorer less happy. Does this mean that low-income countries should give up pro-growth policy? This would be surprising policy advice. Indeed, if relative concerns are important, many may well find it strange to recommend that low-income countries should remain at their current low rank in the concert of nations. Even if income comparisons lead to a vain zero-sum rat race between countries, it is not clear that not competing is an avenue for happiness.

3.3.2 Adaptation

Adaptation and the associated "hedonic treadmill" is another classic explanation of the Easterlin paradox. Habituation effects destroy the welfare benefit of growth. This is because of the deleterious role of aspirations: "*Material aspirations increase commensurately with income, and as a result, one gets no nearer to or farther away from the attainment of one's material goals, and well-being is unchanged*" (Easterlin, 2003).

Adaptation is a central issue in the social sciences: to what extent do we get used to any specific life situation? The psychological basis of adaptation is that judgements of current situations depend on the experience of similar situations in the past, so that higher levels of past experience may offset higher current levels of these phenomena due to changing expectations (see Kahneman and Tversky, 1979). Some psychologists draw a parallel between the homeostasis that leads us to hold body temperature steady and homeostasis in subjective well-being (Cummins, 2003), which latter is argued to hold well-being at some constant individual-specific set-point (argued to be between 60 and 80 on a standardized 0-100 scale, with an average figure of 75). This may be partly biologically determined, underlying a potential role of genetic factors. In any case, the key element is that, although positive and negative events will have short-run effects on well-being, in the longer-run most individuals will return to their set-point level.

Although initially partisans of the adaptation hypothesis, Fujita and Diener (2005) note that in seventeen years of GSOEP data, around one quarter of people changed well-being significantly from the first five to the last five years. Diener *et al.* (2006) propose five significant revisions to hedonic-treadmill theory: 1) individuals' set-points are not hedonically neutral; 2) individuals have different set-points; 3) a single person can have multiple set-points depending on the components of happiness (emotions, life satisfaction); 4) well-being set-points can change under some conditions; and 5) individuals differ in their adaptation to events.

In the context of the Easterlin paradox, we are particularly interested in adaptation to income. With respect to equation (1) above, we again introduce an additional income term into the utility function; however, this time the newcomer is not the income of others or expectations, but rather the income that the individual themselves had earned in the past. Individual well-being is thus still subject to income comparisons, but here the comparisons are within subject, to use the psychological term. Those who have earned more in the past are less satisfied with any given level of income today.

While in theory any past income level could negatively affect well-being today, in practice empirical work has appealed to the income that the individual received one year ago (in panel terms, this is the income that the individual reported in the previous wave, as most panels are carried out on a yearly basis).

(2) $U_{it} = U(y_{it}, y_{it}\text{-}1)$

This kind of utility function implies that any attempt to raise happiness *via* higher income is potentially subject to debate. If the effect (negative) of past income, *via* habituation, is strong enough then income will have no long-lasting well-being effect, at both the individual and the societal level.

Evidence in developed countries

Perhaps the best-cited piece of work in the domain of adaptation to income is that of Brickman *et al.* (1978), who show that a very small sample (22) of American lottery winners report no higher life satisfaction than a control group. The authors' interpretation of this finding is in terms of adaptation to higher income. Much as this paper has been cited, it does not necessarily tell a clean story. Two points of note in this respect are that the winners were actually more satisfied than non-winners, but the small sample size did not yield a significant difference. Further, the analysis is cross-section, rather than panel. As such, it could well be the case that the lottery winners were less happy to start with, before they won. As such, they would have experienced an increase in well-being on winning the lottery, but this would not have been visible in the cross-section analysis.

An early piece of evidence that does appeal to explicit information on income changes is Inglehart and Rabier (1986), who use pooled Eurobarometer data from ten Western European countries between 1973 and 1983 to show that well-being scores are essentially unrelated to current income, but are positively correlated with the change in financial position over the past twelve months. They conclude that aspirations adapt to circumstances, such that, in the long run, stable characteristics do not affect well-being.

More recently, Clark (1999) used two waves of the British Household Panel Survey (BHPS) data to look at the relationship between job satisfaction and current and past labour income. Considering those who stay in the same firm in the same position, past income reduces job satisfaction while current income increases it. This is consistent with a utility function that depends on changes in these variables. The data suggest a completely relative function, with job satisfaction depending only on the annual change in the hourly wage. More recent results in German and British panel data are reported by Di Tella *et al.* (2005) and Burchardt (2005), respectively. Layard *et al.* (2010) appeal to GSOEP data to show that the long-run effect of a rise in income is smaller than the initial effect.

Instead of using own and past individual income, we can also consider aggregate income. Di Tella *et al.* (2003) examine individual happiness in data covering 18 years across 12 European countries, and argue that some of their results on current and lagged GDP *per capita* show that "*bursts of GDP produce temporarily higher happiness*" (p.817).

The Welfare Function of Income, described above, also produces evidence consistent with adaptation to income. In this context, a common finding is that a one-dollar increase in household income leads to a 60-cent increase (within about two years) in the income that individuals consider to be "excellent", "good", "sufficient", "bad" etc. Hence, 60 percent of the welfare effect of income is dissipated by adaptation.

Evidence in LDCs

Much of the work on adaptation to income changes has appealed to panel data to follow individual well-being over time as their income moves around. While there is now a thriving literature looking at adaptation in this way in rich countries, there is at the same time an almost total lack of evidence in poorer countries, undoubtedly due to the lack of panel data in the latter.

Knight and Gunatilaka (2009) is an exception. The work here appeals to data from a household survey for rural China. The survey includes information on life and income satisfaction, but also the minimum income that respondents consider necessary to sustain the household for a year. This latter measure, sometimes known as the Minimum Income Question, was introduced in Goedhart *et al.* (1977). Knight and Gunatilaka consider the answer as a measure of income aspirations. These aspirations are found to be positively correlated with actual income, so that the more individuals earn, the greater the income level they consider as the minimum necessary. Subjective well-being is positively correlated with own income, but negatively correlated with aspiration income. As such, the results are consistent with at least partial adaptation to income in China.[60]

Barr and Clark (2010) analyse South African data, and consider the levels of income that individuals say are necessary to get by, and to live well. In a regression analysis, these are shown to be positively correlated with own income and with reference group income (geographically defined). This is again consistent with a certain amount of adaptation. Along the same lines, Herrera *et al.* (2006) provide a comparative analysis of survey data in Peru and Madagascar. A three-level satisfaction with standard of living variable is shown to be positively correlated with own income, but negatively correlated with average neighbourhood income and the minimum amount the individual thinks is necessary to get by. In turn, this latter minimum amount is positively correlated with own income, suggesting the existence of a ratchet effect whereby higher income increases aspirations and reduces satisfaction.

An impressive piece of evidence by Di Tella and MacCulloch (2010, chapter 8) is based on repeated cross-sections. The authors uncover a positive happiness gradient over time in low-income but not high income countries. In the latter, the level of GDP *per capita* attained in 1960 is sufficient to explain the level of happiness as of 2005. By contrast, in low income countries, both the 1960 level and the later growth in GDP *per capita* exert a statistically significant impact on 2005 subjective well-being. The authors conclude that adaptation is less important in low-income countries: "*The past 45 years of economic growth (from 1960 to 2005) in the rich nations of the world have not brought happiness gains above those that were already in place once the 1960s standard of living had been achieved. However, in the poorest nations, we cannot reject the null hypothesis that the happiness gains they experienced from the past half century of economic growth have been the same as the gains from growth prior to the 1960s. In other words, for these nations, it is still the absolute level of (the logarithm of) income that matters for happiness*" (2010, p.219). This finding with respect to adaptation is thus reminiscent of the concept of threshold effects in the GDP-happiness gradient.

3.3.3. Bounded scales: what exactly is relative?

Is the welfare effect of income purely relative (to other people's income or to one's past level

[60] Castillo's (2010) work mentioned above also shows that income satisfaction in Mexico is positively correlated with the respondent's evaluation of their own current income relative to aspirations. If aspirations rise with own income, then this is also consistent with adaptation.

of income)? Or, on the contrary, could it be the case that happiness measures themselves are relative (to some implicit context)? This question is similar to the distinction made in psychology between the hedonic treadmill (whereby individuals' affect levels gradually adapt back to their initial level following a positive or negative event) and the satisfaction treadmill (in which affect levels do not adapt, but individuals change the way in which they use numbers).

We believe that it is likely that satisfaction judgements expressed on a bounded ordinal scale express relative judgements, *i.e.* the relation between individuals' attainments and the existing possibilities (as represented by the scale). Van Praag (1991), for instance, has illustrated this phenomenon in experimental settings involving bounded scales: subjects tend to divide the total length of the scale into quantiles, equating the higher step with the maximum amount of the proposed magnitude. If this is so, it is not surprising that only a small minority of the population chooses the upper 10th rung on the happiness scale, which is interpreted as "having it all".

Of course, the fact that the happiness scale is interpreted as being context-dependent is difficult to disentangle from happiness itself being context-dependent. However, in order to illustrate the particularity of bounded scales, we separate the quality of life indicators (which are positive correlates of growth) into two groups: the cardinal measures that can be measured on a continuous scale (although often not infinite), such as life expectancy, the percentage of literate population, women's fertility, or the gross enrolment rate in school; and variables that are measured on an ordinal bounded scale, such as happiness, the index of Democracy (Polity IV), the Human Rights index or the Trust variable (see introductory section). Keeping only the countries which were observed for at least ten years in the World Values Survey, and which had experienced an episode of positive growth, we plot the values of these different measures against time. The separate panels of Figure 18 depict these time evolutions in Asian and Western OECD countries.

Two observations are in order. First, objective but ordinal and bounded measures (democracy, human rights) tend to converge to their maximum value as development unfolds *via* GDP growth, whereas the subjective ordinal variables (happiness and trust) remain below the maximum value. Second, the graphs showing average happiness, trust, human rights and democracy tend to be much flatter than those from the cardinal indicators, such as fertility, school-enrolment rates, life expectancy, and infant mortality, which show much clearer trends over time.

In conclusion, we should not therefore necessarily expect bounded ordinal measures to behave like quantitative cardinal measures in the long run. Instead of looking at long-run changes in the average level of subjective well-being (which cannot increase without limit), it is perhaps of more interest to look at the distribution of the answers on the scale proposed. The fact that the variance of SWB tends to fall as GDP grows is on the face of it a promising return to higher GDP for low-income countries.

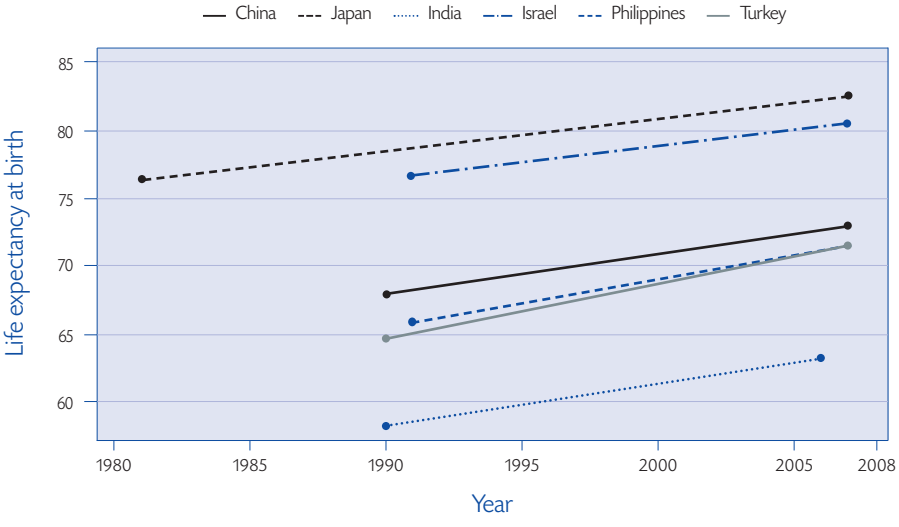
Figure 18

The evolution of cardinal versus *ordinal Quality of Life Indices over a period of growth*

1) Asia:

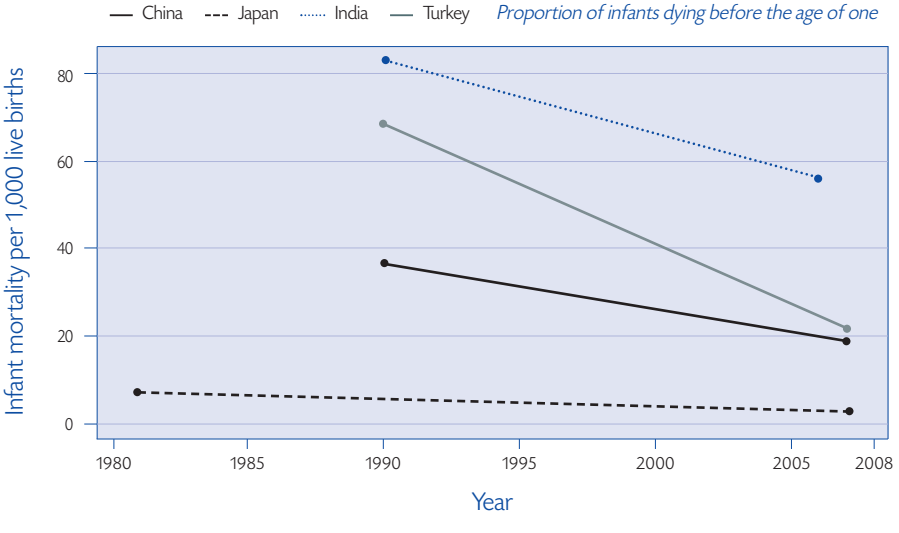
Evolution of life expectancy

Asian countries with a positive growth of GDP in the period



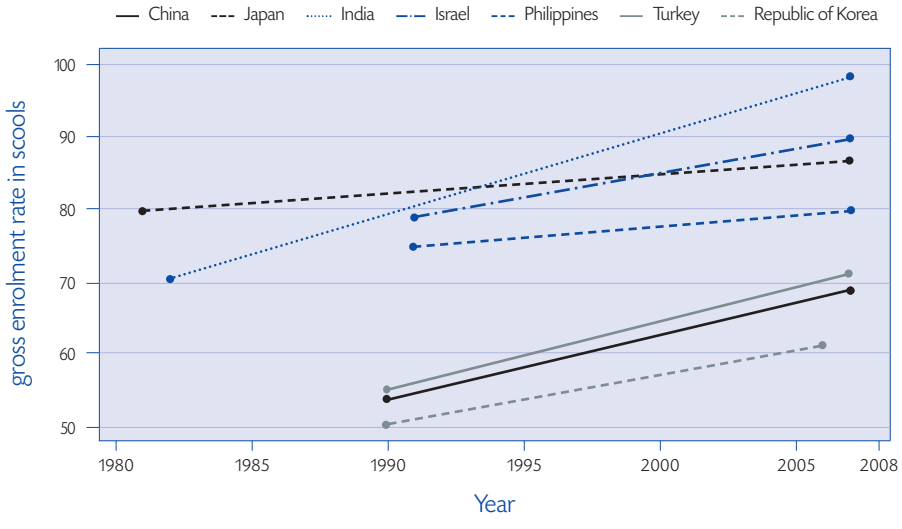
The evolution of infant mortality

Asian countries with a positive growth of GDP in the period



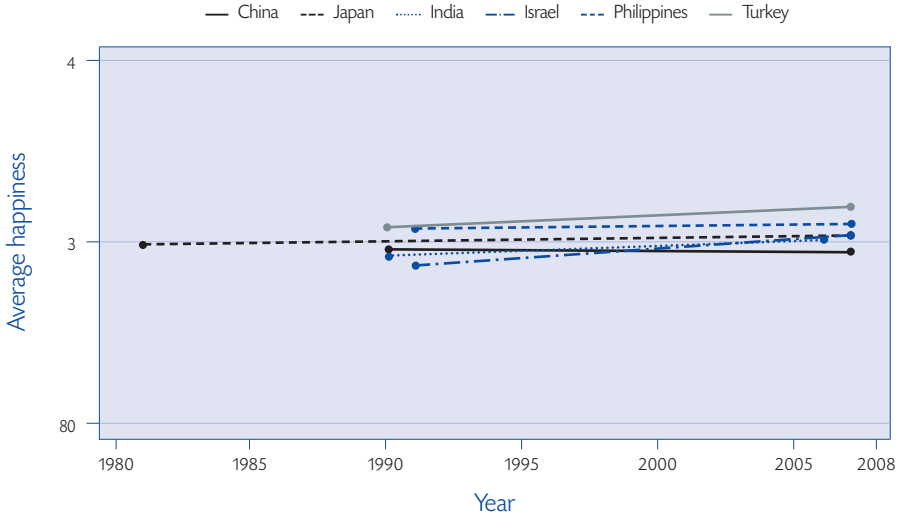
Evolution of the gross enrolment rate in schools

Asian countries with a positive growth of GDP in the period



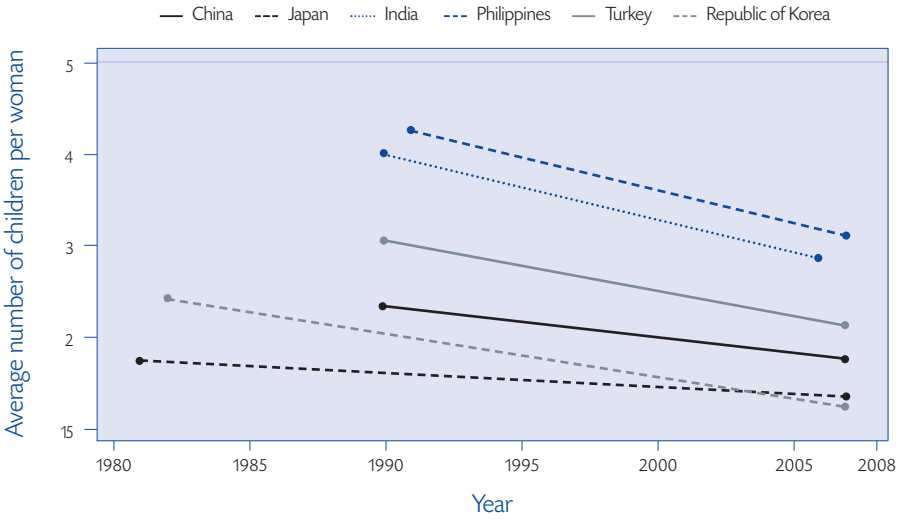
Evolution average happiness

Asian countries with a positive growth of GDP in the period



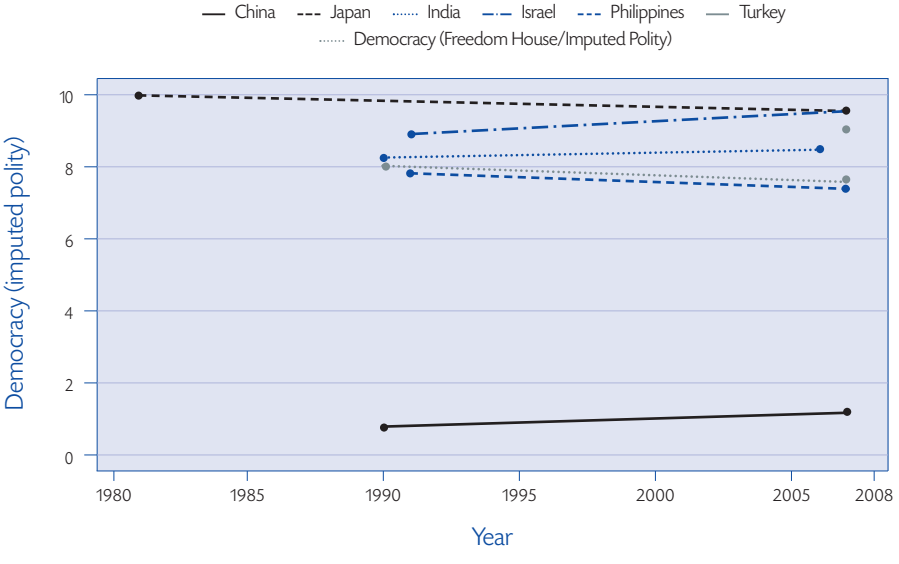
Evolution of fertility

Asian countries with a positive growth of GDP in the period



Evolution of democracy

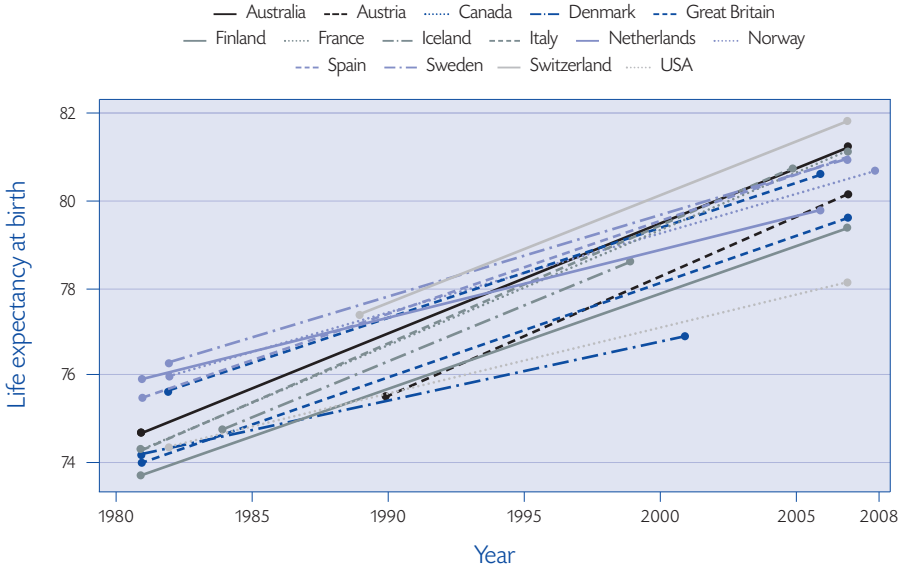
Asian countries with a positive growth of GDP in the period



2) Western countries

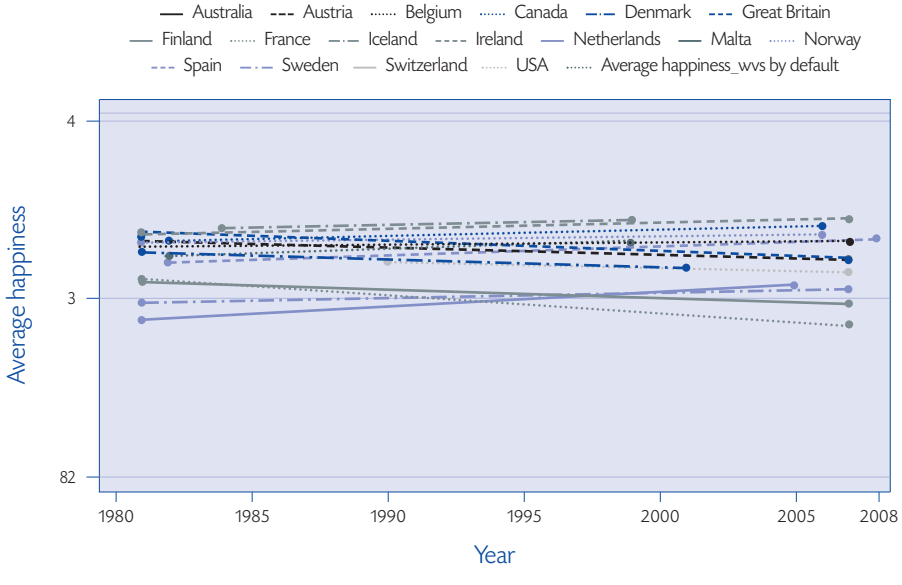
Evolution of life expectancy

Western countries with a positive growth of GDP in the period



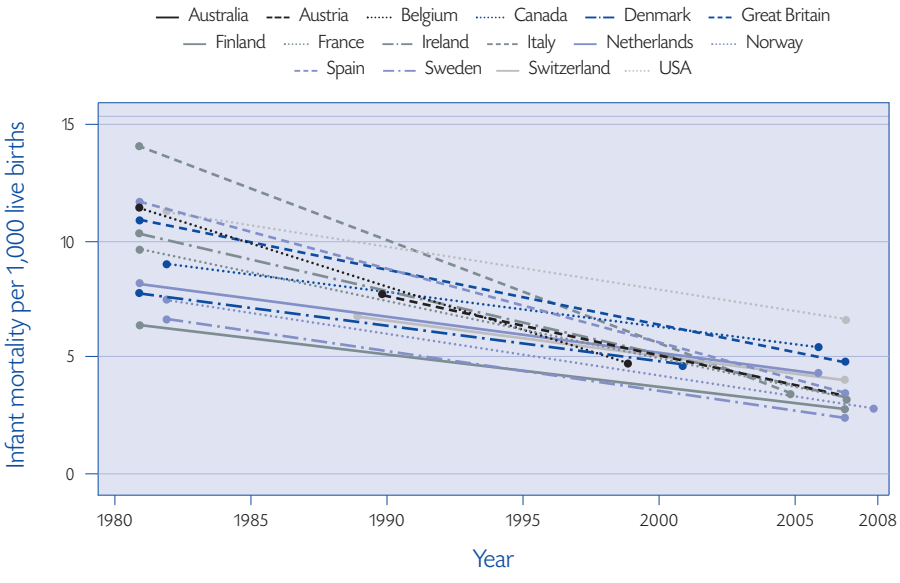
Evolution average happiness

Western countries with a positive growth of GDP in the period



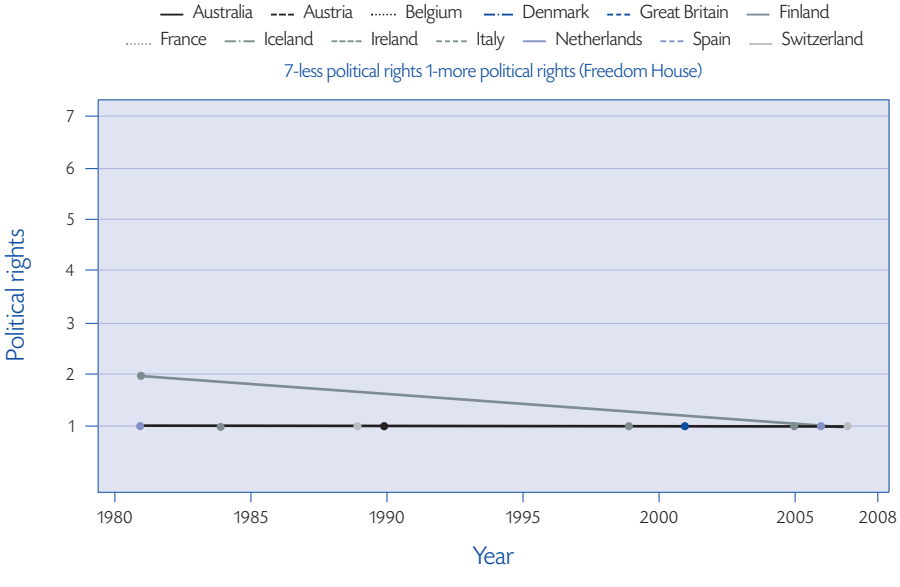
The evolution of infant mortality

Western countries with a positive growth of GDP in the period



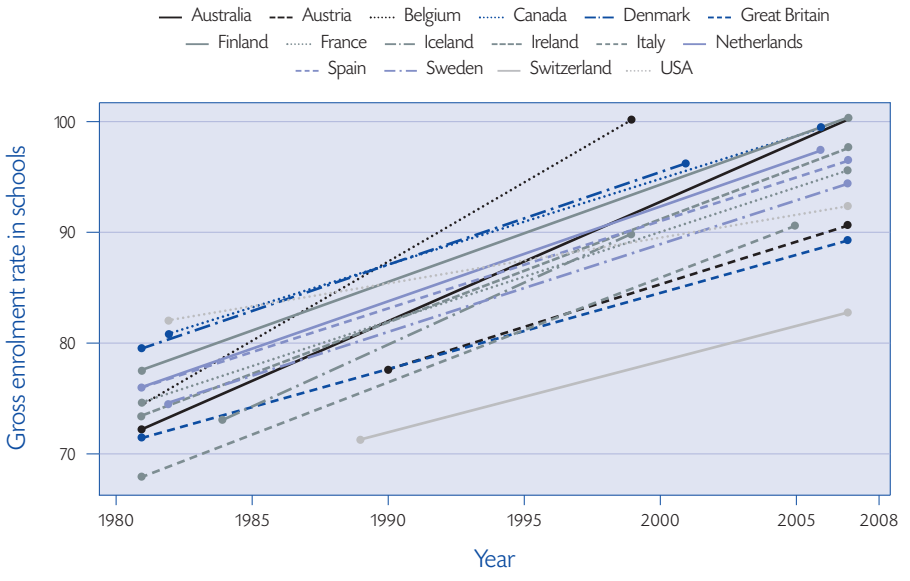
Evolution of political rights

Western countries with a positive growth of GDP in the period



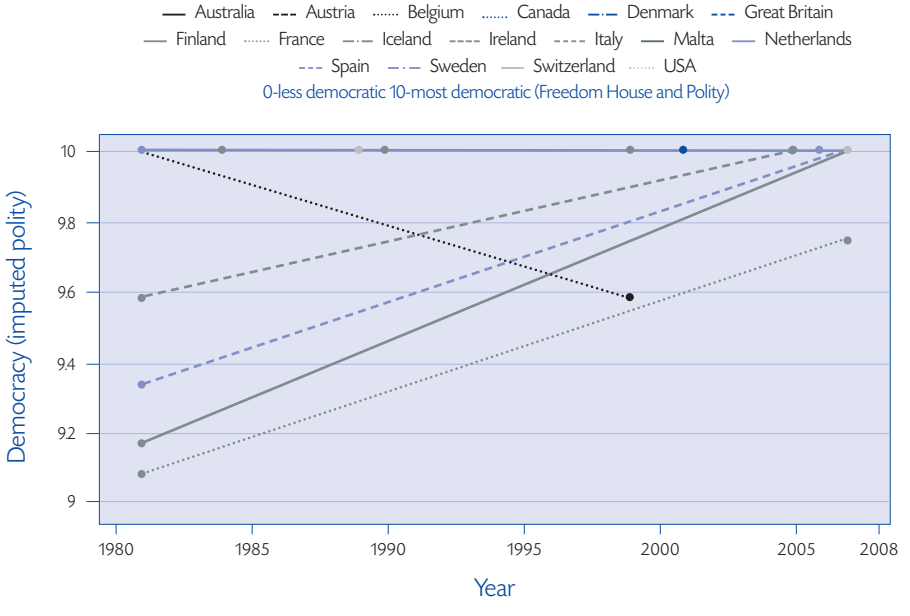
Evolution of the gross enrolment rate in schools

Western countries with a positive growth of GDP in the period



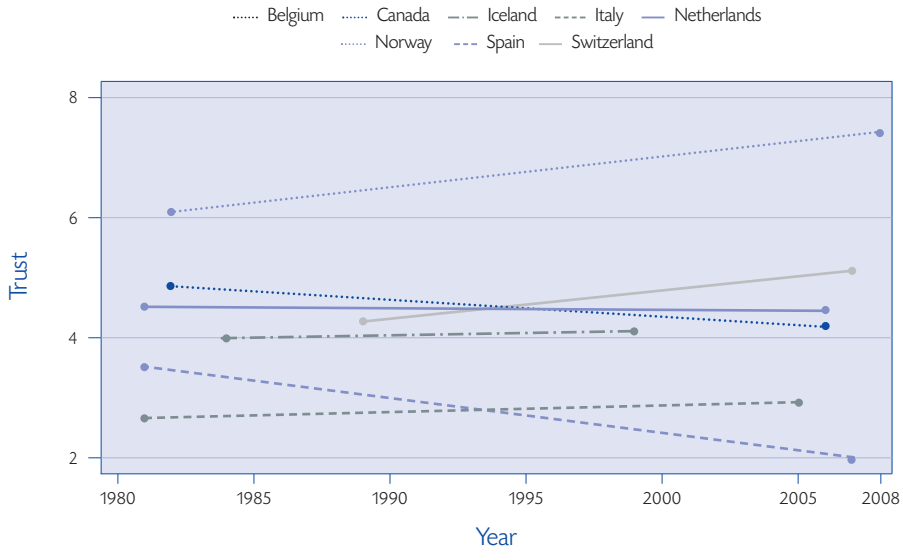
Evolution of democracy

Western countries with a positive growth of GDP in the period



Evolution of trust

Western countries with a positive growth of GDP in the period



Source: WVS (1981-2008).

3.4. Conclusions and take-home messages: how can we use subjective variables in order to understand the GDP-happiness relationship?

The evidence presented in this paper indicates how subjective satisfaction variables can be used in order to measure well-being in developing countries. First of all, subjective well-being measures are particularly well-fitted to capture the multi-dimensional aspect of growth, and can be used to estimate the marginal rates of substitution between different aspects of development that may well have to be traded off against each other, such as higher consumption, greater life expectancy, worsening quality of air, urban congestion, etc. This creates a useful tool for public policy which is aimed at maximizing well-being as countries develop.

Subjective data contain a number of lessons regarding the well-being benefits that growth may confer on developing countries. Cross-sectional data clearly show that income growth yields sizable benefits in terms of self-declared happiness and life satisfaction, although with decreasing marginal returns (*i.e.* the functional form is concave). Within a given country, the richer report higher happiness levels than do the poorer; equally those who live in richer countries are happier than those in poorer countries.

However, the evidence is much less clear-cut regarding long-run changes in well-being, in growing economies. Whether GDP growth yields rising well-being is still hotly debated: essentially, the question is whether the correlation coefficient is "too small to matter". This of course has very important consequences for developing countries, which need to know the potential gains that are associated with growth-oriented policies.

The explanations for the small correlation between income growth and subjective well-being over time appeal to the nature of growth itself, and the way in which humans function psychologically. First, growth may go hand in hand with non-monetary qualitative changes that improve the "quality of life", but may well also be accompanied by unwanted side effects such as pollution, income inequality or stress on the job. Second, greater purchasing power increases individual happiness, but man is a social animal and relative concerns (income comparisons) may well diminish the absolute effect of greater wealth. This is consistent with the positive income-happiness gradient that is regularly observed within countries; it is also consistent with the same

gradient across countries, if income comparisons are global instead of local. A very pessimistic view of growth is then that it may be a zero-sum game, whereby the richer are happier and the poorer less happy, both across populations within a country and across country, but rising income for all may not change the relative income positions. This explains why happiness does not seem to increase with GDP in time-series data. However, even if this is true, many may well find it strange to recommend that low-income countries should remain at their current low rank in the concert of nations. Any single country will always have an incentive to climb up the ranking. The problem is that any gain by one country may well involve losses for other countries, when income is evaluated by comparisons across the globe. Similarly, within a country income growth for one part of the population will benefit them, but may reduce the well-being of others.

An analogous phenomenon is that of adaptation to the standard of living, whereby individuals tend to return to some set-point level of well-being. Growth changes both the environment and aspirations. If both expectations and outcomes increase at the same rate, then individuals will not feel any happier. If they do not realise that their expectations and outcomes tend to move together, individuals will aspire to grow richer, but doing so will not increase their happiness as soon as their expectations catch up with their outcomes. This might be an illusion, as suggested by Easterlin, but can also be seen as some kind of hard-wired mechanism, built into human beings by evolution, to ensure that they keep trying to improve their lot (Rayo and Becker, 2007).

One crucial question in this literature is the relative importance of absolute *versus* relative income concerns. Is the welfare effect of income entirely relative? And is the relative/absolute proportion the same in developing and developed countries? Empirical evidence on the extent of income comparisons is much scarcer in developing countries. The evidence that we do have so far contains two important lessons: income comparisons do seem to affect subjective well-being even in very poor countries; however, adaptation may be more of a rich country phenomenon.

Finally, growth and development do not just concern quantitative increases in consumption, production and the accumulation of capital. They also involve the qualitative transformation of political governance and market development. These qualitative and quantitative processes likely involve take-offs and thresholds. Regime change is an important dimension of these non-linear changes. It is striking that such regime changes are visible in subjective satisfaction measures. The case of transition countries is particularly impressive in this respect: average life satisfaction scores closely mirror changes in GDP for about the first ten years of the transition process, until the regime becomes more stable. By way of contrast, in given stable regimes, such as France, we no longer find any relationship between GDP growth and life satisfaction changes. Our interpretation is that once it becomes stable, the regime becomes the population's frame of reference.

While it is not easy to find large welfare benefits of growth using subjective well-being, there is nonetheless an interesting finding concerning the level and distribution of subjective well-being depending on the country's level of development. The stylized facts are as follows: (i) average SWB rises with GDP *per capita*, but (ii) the standard deviation of SWB

falls with GDP *per capita*. As such, (iii) there is a strong negative relationship between the average and standard deviation of SWB within a country. Consequently, GDP growth reduces the inequality in subjective well-being. This is certainly a desirable outcome. If individuals are risk averse, then behind the veil of ignorance they would prefer a society in which well-being is more equally distributed, *ceteris paribus*.

The recourse to subjective measures of well-being is particularly welcome for assessing social phenomena that are not measurable using the standard approach of revealed preference. Whenever social interactions, social preferences or externalities are involved, it becomes more difficult to trace out the link from individual preferences to individual actions. There is no price one can pay to buy less inflation, unemployment or income inequality.

However, subjective variables should be used as a complement to action-revealed preferences, rather than as a replacement. When people clearly vote with their feet, it is difficult to dismiss their actions on the ground that the message is not confirmed in subjective data. With respect to growth and well-being, as long as international migrations remain clearly unidirectional, from low-to high-income countries, it would appear extremely difficult to argue that GDP growth, in the mind of less-developed countries, does not bring higher well-being. The revealed preferences here are consistent with the cross-sectional evidence of a positive income-well-being gradient.

Our stand is that the dynamic evidence based on subjective well-being is much less solid than the cross-sectional and panel evidence, based on individual data. This is because cross-country time-series comparisons are based on aggregate measures, which have lower variance and are less powerful in terms of statistical inference. Moreover, it is possible that the satisfaction judgements expressed on a bounded scale yield relative judgements by their very nature, due to the relation between outcomes and the set of possibilities (represented by the bounded scale). In this case, it is to be expected that only a small minority of individuals choose the 10th rung on the scale, which is interpreted as "having it all". *De facto*, quantitative variables, such as fertility, life expectancy or literacy, exhibit much clearer trends over time than do these bounded-scale qualitative variables, such as governance indicators.

The relationship between income growth and well-being is still the object of ongoing debates that would undoubtedly be better illuminated by the development of panel surveys of the populations of low-income countries.

Appendix 1.

Table 7 *Descriptive statistics of the variables from the WVS database*

Variable	N°. Countries	N°. Years	Mean value	Std. Dev.	Min.	Max.	Source	First year	Last year
Average happiness	105	368	3.02	0.27	2.06	3.58	WVS	1981	2008
Average satisfaction	98	251	6.70	1.07	3.73	8.49	WVS	1981	2008
N°. of children per woman	96	342	2.12	1.02	1.08	6.79	WDI	1981	2008
Democracy (Freedom House/Imputed Polity)	96	337	8.24	2.34	0.00	10.00	Freedom house	1981	2008
Political Rights	96	337	2.15	1.65	1.00	7.00	Freedom house	1981	2008
GDP growth per year	97	348	3.12	4.82	-14.57	46.50	WDI	1981	2008
GDP *per capita* in constant 2000 dol	97	348	11536.26	11138.87	175.01	43420.52	WDI	1981	2008
Growth GDP *per capita*	97	348	2.32	4.64	-14.57	42.86	WDI	1981	2008
GDP *per capita* in PPP	96	347	16508.83	11547.53	236.94	57034.16	WDI	1981	2008
Gross enrolment rate	95	331	78.87	12.90	32.77	100.00	HDI	1981	2007
Gini index	49	91	39.82	11.80	19.40	60.24	WDI	1989	2007
Life expectancy at birth	96	344	72.89	6.80	42.19	82.51	WDI	1981	2008
Infant mortality rate per 1,000	74	254	14.23	19.52	2.50	120.00	WDI	1981	2008
Average trust	98	251	0.30	0.15	0.03	0.74	WVS	1981	2008

Variables description

All variables are available in the World Data Bank: http://www.worldvaluessurvey.org/.

Happiness: "*If you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole?*" (the question and different response categories are the same in the three studies): 1. Not at all happy; 2. Not very happy; 3. Fairly happy; and 4. Very happy.

Life satisfaction: "*All things considered, how satisfied are you with your life as a whole these days?*" The response categories go from 1 (dissatisfied) to 10 (very satisfied).

Trust: "*Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?*", Answers: 1. Most people can be trusted; 0. Can't be too careful.

Fertility rate: this measure represents the number of children that would be born to a woman were she to live to the end of her childbearing years and bear children in accordance with the current age-specific fertility rates.

GDP growth: annual percentage growth rate of GDP at market prices in constant local currency. Aggregate figures are based on constant 2000 US dollars.

GDP *per capita* in 2000 dollars: GDP *per capita* is gross domestic product divided by midyear population. Data are in constant US dollars.

Gini index: the Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality.

Life expectancy at birth: life expectancy at birth indicates the number of years a newborn infant would live were prevailing patterns of mortality at the time of its birth to stay the same throughout their life.

Infant mortality rate (per 1,000 under one): the number of infants dying before reaching age one, per 1,000 live births in a given year.

Gross enrolment rate in %: enrolment in primary, second and tertiary education.

Adult literacy rate in %.

Freedom House: (http://www.freedomhouse.org): political rights that enable people to participate freely in the political process, including the right to vote freely for distinct alternatives in legitimate elections, compete for public office, join political parties and organizations, and elect representatives who have a decisive impact on public policies and are accountable to the electorate. The specific list of rights considered varies over the years. Countries are graded between 1 (most free) and 7 (least free).

Democracy: average of Freedom House and Polity, transformed to a scale 0-10, where 0 is least democratic and 10 most democratic (http://www.govindicators.org).

Appendix 2.

The income-happiness nexus: sources and estimates: a summary

Subjective well-being measures

Happiness: *if you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole: Not at all happy; Not very happy; Fairly happy; Very happy?*

Life satisfaction: *All things considered, how satisfied are you with your life as a whole these days? 1 (dissatisfied) – 10 (very satisfied).*

1) The static relationship between individual income and individual happiness

Consensus: higher income ? → higher happiness. In a country, richer individuals are happier than poorer individuals.

Nationally representative household surveys. Individual-level analysis. Within-country cross-section estimates.

Western developed countries

German Socio-Economic Panel (GSOEP), British Household Panel Survey (BHPS), Swiss household panel, Australian household survey (HILDA), General Social Survey (America), Japanese household survey, data from Netherlands, Denmark, etc.

European Values Survey (EVS), European Social Survey (ESS), Eurobarometer.

Transition countries

Albania, Belarus, Bulgaria, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Russia, Ukraine, etc.

Life in Transition Survey (LITS, 2006), European Social Survey, European Values Survey.

Asian household surveys

China, India, Shanghai.

African and Middle-East national household surveys

Argentina, Brazil, Chile, Ethiopia, Mexico, Mexico, Nigeria, Peru, South Korea, South Africa (SALDRU), Tanzania, Turkey, Venezuela.

International surveys

- World Values Survey (WVS, 1981-2008, 5 waves, 105 countries).
- International Social Survey Program (ISSP, 101 countries).
- Gallup World Poll (2006, 105 countries).
- Latinobarometer (18 countries).
- European Social Survey (25 countries).
- European Values Survey.

2) The dynamic relationship between individual income and individual happiness

Within-country estimates. Individual-level panel data analysis.

Consensus: higher income → higher happiness. Individuals become happier as they grow richer.

Individual Panel Data in Developed Countries

GSOEP, BHPS, HILDA, data from Netherlands and Denmark.

Individual Panel data in LDCs

RLMS (Russia), ULMS (Ukraine), Peru, LSMS (Tajikistan).

3) The static relationship between national income and average happiness

Aggregate measures, cross-country estimates.

Consensus: higher income → higher happiness. Individuals living in richer countries are happier than those living in poorer countries.

4) The dynamic relationship between national income and average happiness

Aggregate measures, cross-country estimates.

No consensus. Divergent findings.

Income growth does not increase happiness over time

- Easterlin (2005*a*), Easterlin and Sawangfa. (2005, 2009), Easterlin and Angelescu (2007), Easterlin (2009).
- Layard, Brockmann (2003,2005).
- Delhey, Welzel, Yuan (2009).

Income growth does increase happiness over time

- Stevenson and Wolfers (2008).
- Deaton (2008), Gallup (2006).
- Helliwell (2002).
- Blanchflower (2008).

Income growth does increase happiness over time but not always and weakly

- Hagerty and Veenhoven (2000, 2003, 2006), WVS (positive and statistically significant coefficient, but not in all countries).
- Inglehart, Peterson and Welzel (2008): WVS, BHPS, GSS (positive and statistically significant coefficient, but not in all countries), Kenny (2005), *idem.*
- Layard, Mayraz and Nickell (2010) (positive coefficient but not always statistically significant).
- Oswald (1997) in GSS and Eurobarometer survey series, positive coefficient but not always statistically significant.
- Di Tella and MacCulloch (2008): positive coefficient but low statistical significance.

References

Akay, A. and P. Martinsson (2008), "Does Relative Income Matter for the Very Poor? Evidence from Rural Ethiopia", IZA, Discussion Paper No. 3812.

Akay, A., P. Martinsson and H. Medhin (2009), "Does Positional Concern Matter in Poor Societies? Evidence from a Survey Experiment in Rural Ethiopia". IZA, Discussion Paper No. 4354.

Alesina, A. and P. Giuliano (2008), Preferences for Redistribution, Harvard University and UCLA mimeo.

Alpizar, F., F. Carlsson and O. Johansson-Stenman (2005), "How much do we care about absolute *versus* relative income and consumption?", *Journal of Economic Behavior and Organization*, 56: 405-421.

Andersson, F. (2006), "Is Concern for Relative Consumption a Function of Relative Consumption?", Göteborg University, mimeo.

Appleton, S. and L. Song (2008), "Life Satisfaction in Urban China: Components and Determinants", *World Development*, 36: 2325-2340.

Barr, A. and D. Clark (2010), "Do the Poor Adapt to Low Income, Minimal Education and Ill-Health?", *Journal of African Economies*, 19: 257-293.

Becker, G.S., T. J. Philipson and R.R. Soares (2005), "The quantity and quality of life and the evolution of world inequality", *American Economic Review*, 95: 277-291.

Blanchflower D.G. (2008), "International evidence on well-being", IZA Discussion Paper No. 3354.

Blanchflower D.G. and A. Oswald (2004), "Well-being over time in Britain and the USA", *Journal of Public Economics*, 88: 1359-1386.

Blanchflower D.G. and A. Oswald (2005), "Happiness and the Human development index: the paradox of Australia", *Australian Economic Review*, 38: 307-318.

Bookwalter, J. and D. Dalenberg (2010), "Relative to What or Whom? The Importance of Norms and Relative Standing to Well-Being in South Africa", *World Development*, 38: 345-355.

Brickman, P., D. Coates and R. Janoff-Bulman (1978), "Lottery winners and accident victims: is happiness relative?", *Journal of Personality and Social Psychology*, 36: 917-927.

Brockmann, H., J. Delhey, C.Welzel and H. Yuan (2009), "The China Puzzle: Falling Happiness in a Rising Economy", *Journal of Happiness Studies*, 10: 387-405.

Brown, G., J. Gardner, A. J. Oswald and J. Qian (2008), "Does Wage Rank Affect Employees' Wellbeing?", *Industrial Relations*, 47: 355 - 389.

Brown, P., E. Bulte and X. Zhang (2010), "Positional Spending and Status Seeking in Rural China", *Journal of Development Economics*, 96 (1).

Burchardt, T. (2005), "Are One Man's Rags Another Man's Riches? Identifying Adaptive Preferences Using Panel Data", *Social Indicators Research*, 74: 57-102.

Card, D., A. Mas, E. Moretti and E. Saez (2010), "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction", NBER Working Paper No. 16396.

Carlsson, F., G. Gupta and O. Johansson-Stenman (2009), "Keeping up with the Vaishyas Caste and relative standing in India", *Oxford Economic Papers*, 61: 52-73.

Castilla, C. (2010), "Subjective poverty and reference dependence", UNU-WIDER Working Paper No. 2010/76.

Clark, A.E. (1996), "L'utilité est-elle relative? Analyse à l'aide de données sur les ménages", *Economie et Prévision*, 121: 151-164.

Clark, A.E. (1999), "Are Wages Habit-Forming? Evidence from Micro Data", *Journal of Economic Behavior and Organization*, 39: 179-200.

Clark, A.E. (2001), "What Really Matters in a Job? Hedonic Measurement Using Quit Data", *Labour Economics*, 8: 223-242.

Clark, A.E. (2003), "Unemployment as a Social Norm: Psychological Evidence from Panel Data", *Journal of Labor Economics*, 21: 323-351.

Clark, A.E., P. Frijters and M. Shields (2008), "Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles", *Journal of Economic Literature*, 46: 95-144.

Clark, A.E., Y. Georgellis and P. Sanfey (1998), "Job Satisfaction, Wage Changes and Quits: Evidence from Germany", *Research in Labor Economics*, 17: 95-121.

Clark, A.E. and J.A.V. Fischer, (2009), in OECD, 2009, Subjective well-being and social policy.

Clark, A.E., A. Knabe and S. Rätzel (2010), "Boon or Bane? Others' Unemployment, Well-being and Job Insecurity", *Labour Economics*, 17: 52-61.

Clark, A.E., N. Kristensen and N. Westergård-Nielsen (2009), "Economic Satisfaction and Income Rank in Small Neighbourhoods", *Journal of the European Economic Association*, 7: 519-527.

Clark, A.E. and A.J. Oswald (1996), "Satisfaction and Comparison Income", *Journal of Public Economics*, 61: 359-81.

Clark, A.E. and C. Senik (2010), "Who compares to whom? The anatomy of income comparisons in Europe", *Economic Journal*, 120: 573-594.

Cohen, S., W. Doyle, R. Turner, C. Alper and D. Skoner (2003), "Emotional style and susceptibility to the common cold", *Psychosomatic Medicine*, 65: 652-657.

Cojocaru, A. (2010), "On relative status inference when reference groups are unobserved: a cautionary tale", University of Maryland, mimeo.

Corazzini L., L. Esposito and F. Majorano, "Reign in Hell or Serve in Heaven? A Cross-country Journey into the Relative *vs* Absolute Perceptions of Well-being", University of Padua, mimeo.

Cummins, R. (2003), "Normative life satisfaction: measurement issues and a homeostatic model", *Social Indicators Research*, 64: 225-256.

Danner, D., D. Snowdon and W. Friesen (2001), "Positive Emotions in Early Life and Longevity: Findings from the Nun Study", *Journal of Personality and Social Psychology*, 80: 804-813.

Davidson, R. (2004), "Well-being and affective style: neural substrates and biobehavioural correlates", *Philosophical Transactions of the Royal Society*, B, 359: 1395-1411.

Deaton, A. (2008), "Income, Health and Well-Being around the World: Evidence from the Gallup World Poll", *Journal of Economic Perspectives*, 22: 53-72.

Di Tella, R., J. Haisken-De New and R. MacCulloch (2007), "Happiness Adaptation to Income and to Status", Harvard Business School, mimeo.

Di Tella, R. and R. MacCulloch (2006), "Some Uses of Happiness Data in Economics", *Journal of Economic Perspectives*, 20: 25-46.

Di Tella, R. and R. MacCulloch (2008), "Gross National Happiness as an Answer to the Easterlin Paradox?", *Journal of Development Economics*, 86, 22-42.

Di Tella, R. and R. MacCulloch, (2010),"Happiness Adaptation to Income Beyond 'Basic Needs'", in E. Diener, J. Helliwell, and D. Kahneman (eds), *International Differences in Well-Being, Oxford, Oxford University Press.*

Di Tella, R., R.J. MacCulloch and A.J. Oswald (2001), "Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness", *American Economic Review*, 91, 335-341.

Di Tella, R., R.J. MacCulloch and A.J. Oswald (2003), "The Macroeconomics of Happiness", *Review of Economics and Statistics*, 85: 809-827.

Diener E., J. Helliwell and D. Kahneman (eds) (2009), *International Differences in Well-being*, Princeton University Press, Princeton.

Diener, E. and R.E. Lucas (1999), "Personality and subjective well-being", in D. Kahneman, E. Diener and N. Schwarz (eds), *Foundations of Hedonic Psychology: Scientific Perspectives on Enjoyment and Suffering*, Chapter 11, Russell Sage Foundation, New York.

Diener, E., R. Lucas and C. Scollon (2006), "Beyond the Hedonic Treadmill", *American Psychologist*, 61: 305-314.

Easterlin, R. (1974), Does Economic Growth Improve the Human Lot?, *Nations and Households in Economic Growth*, P.A. David and W.B. Melvin (eds), Stanford University Press, Palo Alto: 89-125.

Easterlin, R.A. (1995), "Will Raising the Incomes of All Increase the Happiness of All?", *Journal of Economic Behavior and Organization*, 27: 35-47.

Easterlin R. (2000), "The worldwide standard of living since 1800", *Journal of Economic Perspectives*, 14: 7-26.

Easterlin, R.A. (2001), "Income and Happiness: Towards a Unified Theory", *Economic Journal*, 111: 465-484.

Easterlin, R. (2003), "Explaining Happiness", *Proceedings of the National Academy of Science*, 100: 11176-11183.

Easterlin, R. (2005*a*), "Diminishing Marginal Utility of Income? Caveat Emptor", *Social Indicators Research* 70: 243-255.

Easterlin, R. (2005*b*), "Feeding the Illusion of Growth and Happiness: A Reply to Hagerty and Veenhoven", *Social Indicators Research*, 74: 429-443.

Easterlin R. and L. Angelescu (2009), "Happiness and growth the world over: time series evidence on the Happiness-income paradox", IZA Discussion Paper N°. 4060.

Easterlin R. and L. Angelescu (2010), "Modern Economic Growth: Cross Sectional and Time Series Evidence" in K. C. Land, (ed.), *Handbook of Social Indicators and Quality of Life Research*, forthcoming, Springer, New York and London.

Easterlin R. and O. Sawangfa (2010), "Happiness and Growth: Does the Cross Section Predict Time Trends? Evidence from Developing Countries", in E. Diener, J. Helliwell, and D. Kahneman, (eds)„ International Differences in Well-Being, Princeton, Princeton University Press, NJ, chapter 7: 162-212.

Easterlin R. and A. Zimmermann (2006), "Life satisfaction and economic outcomes in Germany in pre and post unification", IZA Discussion Paper N°. 2494.

Easterlin R. and A. Zimmermann (2009), "Lost in Transition: Life Satisfaction on the Road to Capitalism", *Journal of Economic Behavior and Organization*, 71: 130-145.

Fafchamps, M. and F. Shilpi (2008), "Subjective welfare, isolation, and relative consumption", *Journal of Development Economics*, 86: 43-60.

Ferrer-i-Carbonell, A. (2005), "Income and well-being: an empirical analysis of the comparison income effect", *Journal of Public Economics*, 89: 997-1019.

Ferrer-i-Carbonell, A. and P. Frijters (2004), "How important is methodology for the estimates of the determinants of happiness?", *Economic Journal*, 114: 641-659.

Fließbach, K., B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. Elger and A. Falk (2007), "Social comparison affects reward-related brain activity in the human ventral striatum", *Science*, 318: 1305-1308.

Freeman, R. (1978), "Job satisfaction as an economic variable", *American Economic Review*, 68: 135-141.

Frey, B.S. and A. Stutzer (2000), "Happiness, Economy and Institutions", *Economic Journal*, 110: 918-938.

Frey, B.S. and A. Stutzer (2007), "Should National Happiness be Maximized?", University of Zurich, mimeo.

Fujita, F. and E. Diener (2005), "Life Satisfaction Set Point: Stability and Change", Journal of *Personality and Social Psychology*, 88: 158-164.

Gao, W. and R. Smyth (2010), "Job satisfaction and relative income in economic transition: status or signal? The case of urban China", *China Economic Review*, 21: 442-455.

Gardner, J. and A.J. Oswald (2006), "Do Divorcing Couples Become Happier by Breaking Up?", *Journal of the Royal Statistical Society*, 169: 319-336.

Graham, C. and A. Felton (2006), "Inequality and happiness: Insights from Latin America", *Journal of Economic Inequality*, 4: 107-122.

Goedhart, T., V. Halberstadt, A. Kapteyn and B.M. Van Praag (1977), "The poverty line: Concept and measurement", *Journal of Human Resources*, 12: 03-520.

Guillen-Royo, M. (2010), "Reference group consumption and the subjective well-being of the poor in Peru", *Journal of Economic Psychology*, forthcoming.

Guriev, S. and E. Zhuravskaya (2009), "(Un)Happiness in Transition", *Journal of Economic Perspectives*, 23: 143-168.

Guven, C., C. Senik and H. Stichnoth (2010), "You Can't Be Happier than Your Wife. Happiness Gaps and Divorce", IZA, Discussion Paper No. 4599.

Hagerty, M.R. and R. Veenhoven (2000), "Wealth and Happiness Revisited - Growing Wealth of Nations *Does* Go with Greater Happiness", University of California-Davis, mimeo.

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

Herrera, J., M. Razafindrakoto and F. Roubaud (2006), "The determinants of subjective poverty: A comparative analysis in Madagascar and Peru", DIAL Working Paper 2006-01.

Hung-Lin, T. and S. Chiu "Income growth or equal income distribution? Implications of the relative income effect on happiness", Department of Economics, Soochow University of Taiwan Working Paper.

Inglehart, R., R. Foa, C. Peterson and C. Welzel (2008), "Development, Freedom, and Rising Happiness: A Global Perspective (1981-2007)", *Perspectives on Psychological Science*, 3: 264-285.

Inglehart, R. and J.R. Rabier (1986), "Aspirations Adapt to Situations - But Why Are the Belgians so Much Happier than the French? A Cross-Cultural Analysis of the Subjective Quality of Life", in F.M. Andrews (ed.), *Research on the Quality of Life*, Institute for Social Research University of Michigan, Ann Arbor.

Inter-American Development Bank (2008), *Beyond Facts: Understanding Quality of Life*, Harvard University Press, Cambridge, Mass.

Johannsson-Stenman, O., F. Carlsson and D. Daruvala (2002), "Measuring Future Grandparents' Preferences for Equality and Relative Standing", *Economic Journal*, 112: 362-383.

Kahneman, D. and A. Deaton, 2010. "High income improves evaluation of life but not emotional well-being", PNAS, Princeton Center for Health and Well-Being.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47: 263-291.

Kenny, C. (2005), "Does development make you happy? Subjective well-being and economic growth in developing countries", *Social Indicators Research*, 73: 199-219.

Kingdon, G. and J. Knight (2007), "Community, Comparisons and Subjective Well-being in a Divided Society", *Journal of Economic Behavior and Organization*, 64: 69-90.

Knight, J. and R. Gunatilaka (2009), "Income, Aspirations and the Hedonic Treadmill in a Poor Society", University of Oxford Discussion Paper No.468.

Knight, J. and R. Gunatilaka (2010*a*), "Great Expectations? The Subjective Well-being of Rural–Urban Migrants in China", *World Development*, 38: 113-124.

Knight, J. and R. Gunatilaka (2010*b*), "The rural-urban divide in China: Income but not happiness?", *Journal of Development Studies*, 46: 506-29.

Knight, J., L. Song and R. Gunatilaka (2007), "Subjective well-being and its determinants in rural China", *China Economic Review*, 20: 635-649.

Kristensen, N. and N. Westergaard-Nielsen (2006), "Job satisfaction and quits - Which job characteristics matters most?", *Danish Economic Journal*, 144: 230-248.

Kuegler, A. (2009), "A Curse of Comparison? Evidence on Reference Groups for Relative Income Concerns", World Bank, Policy Research Working Paper Series N°. 4820.

Layard, R. (2003), "Happiness: Has Social Science A Clue?", LSE, mimeo.

Layard, R. (2005), *Happiness: Lessons from a New Science*, Penguin, London.

Layard, R., G. Mayraz and S. Nickell (2008), "The Marginal Utility of Income", *Journal of Public Economics*, 92: 1846-1857.

Layard, R., G. Mayraz and S. Nickell (2010), "Does Relative Income Matter? Are the Critics Right?", in E. Diener, J.F. Helliwell, and D. Kahneman (eds), *International Differences in Well-Being*, Oxford University Press, New York.

Luttmer, E. (2005), "Neighbors as Negatives: Relative Earnings and Well-Being", *Quarterly Journal of Economics*, 120 : 963-1002.

McBride, M. (2001), "Relative-income Effects on Subjective Well-being in the Cross-section", *Journal of Economic Behavior and Organization*, 45: 251-278.

Mishra, V., I. Nielsen and R. Smyth (2010), "Relative Income, Temporary Life Shocks and Subjective Well-being in the Long-run", Monash University, mimeo.

Oswald, A.J. (1997), "Happiness and economic performance", *Economic Journal*, 107: 1815-31.

Palmore, E. (1969), "Predicting Longevity: A Follow-Up Controlling for Age", *Gerontologist*, 9: 247-50.

Pischke, S. (2010), "Money and Happiness: Evidence from the Industry Wage Structure", LSE, mimeo.

Polity IV Project: Political Regime Characteristics and Transitions, 1800-2009, M.G. Marshall, Director, M. G. Marshall and K. Jaggers, Principal Investigators, Center for Systemic Peace and Colorado State University; Ted Robert Gurr, Founder, University of Maryland, http://www.systemicpeace.org/polity/polity4.htm.

Pradhan, M. and M. Ravallion (1998), "Measuring poverty using qualitative perceptions of welfare", World Bank and the University of Amsterdam, Working Paper.

Pritchett, L. and L. Summers (1996), "Wealthier is Healthier", *Journal of Human Resources*, 31: 841-868.

Ravallion, M. and M. Lokshin (2010), "Who cares about relative deprivation?" *Journal of Economic Behavior & Organization*, 73 : 171-185.

Rayo, L. and G.S. Becker (2007), "Evolutionary Efficiency and Happiness", *Journal of Political Economy*, 115: 302-37.

Rojas, M. (2010), "Relative Income and Well-Being in Latin America", Universidad Popular Autónoma del Estado de Puebla, mimeo.

Rojas, M. and E. Jiménez (2007), "Pobreza Subjetive en México: El Papel de las Normas de Evaluación de Ingreso", Universidad Popular Autónoma del Estado de Puebla, mimeo.

Sandvik, E., E. Diener and L. Seidlitz (1993), "Subjective well-being: the convergence and stability of self and non self report measures", *Journal of Personality*, 61: 317-342.

Sen, A. (2001), *Development as Freedom*, Alfred Knopf, New York.

Senik, C. (2004), "When Information Dominates Comparison: A Panel Data Analysis Using Russian Subjective Data", *Journal of Public Economics*, 88 : 2099-2123.

Senik, C. (2005), "Income Distribution and Well-Being: What Can we Learn from Subjective Data?", *Journal of Economic Surveys*, 19: 43-63.

Senik, C. (2008), "Ambition and jealousy. Income interactions in the 'Old' Europe *versus* the 'New' Europe and the United States", *Economica*, 75 : 495-513.

Senik, C. (2009), "Income Distribution and Subjective Happiness: A Survey", OECD Social Employment and Migration Working Papers N°. 96. Published as Income distribution and subjective well-being, in OECD, 2009, Subjective well-being and social policy.

Smyth, R. and X. Qian (2008), "Inequality and Happiness in Urban China", *Economics Bulletin*, 4:1-10.

Solnick, S. and D. Hemenway (2005), "Are Positional Concerns Stronger in Some Domains than in Others?", *American Economic Review*, 95: 147-151.

Stark, O. and J.E. Taylor (1991), "Migration Incentives, Migration Types: The Role of Relative Deprivation", *Economic Journal*, 101: 1163-1178.

Stevenson, B. and J. Wolfers (2008), "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox", *Brookings Papers on Economic Activity*, Spring.

Stutzer, A. (2004), "The Role of Income Aspirations in Individual Happiness", *Journal of Economic Behaviour and Organization*, 54(1): 89-109.

Takahashi, H., M. Kato, M. Matsuura, D. Mobbs, T. Suhara and Y. Okubo (2009), "When your gain is my pain and your pain is my gain: Neural correlates of envy and Schadenfreude", *Science*, 323: 937-939.

Urry, H.L., J.B. Nitschke, I. Dolski, D.C. Jackson, K.M. Dalton, C. J. Mueller, M.A. Rosenkranz, C.D. Ryff, B.H. Singer and R.J. Davidson (2004), "Making a life worth living: Neural correlates of well-being", *Psychological Science*, 15: 367-372.

Van de Stadt, H., A. Kapteyn and S. Van de Geer (1985), "The Relativity of Utility: Evidence from Panel Data", *Review of Economics and Statistics*, 67: 179-187.

Van Praag, B.M. (1971), "The Welfare Function of Income in Belgium: An Empirical Investigation", *European Economic Review*, 2: 337-369.

Van Praag, B.M. (1991), "Ordinal and Cardinal Utility", *Journal of Econometrics*, 50: 69-89.

Van Praag, B. and P. Frijters (1999), "The measurement of welfare and well-being: the Leyden approach", in D. Kahneman, E. Diener and N. Schwarz (eds) *Well-being; The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Vendrik, M. and G. Woltjer (2006), "Happiness and Loss Aversion: When Social Participation Dominates Comparison", IZA, Discussion Paper No. 2218.

Wolfers, J. (2003), "Is Business Cycle Volatility Costly? Evidence from Surveys of Subjective Well-being", *International Finance*, 6: 1-26.

Wolfhard, K. (2010), "Conspicuous Consumption and Race: Evidence from South Africa", Papers on Economics and Evolution 2010-03, Max Planck Institute of Economics, Evolutionary Economics Group.

Zizzo, D.J. and A.J. Oswald (2001), "Are People Willing to Pay to Reduce Others' Incomes?", *Annales d'Economie et de Statistique*, 63-64: 39-65.

Comment: Will GDP Growth Raise Well-Being in Developing Countries?

Pramila Krishnan, Faculty of Economics, University of Cambridge

The answer to this question is Yes: in fact, we have a formula which is life evaluation = 0.25 ln (income).[61] As Clark and Senik (2010) have explained in their admirably detailed presentation of the evidence, this means that increases in income increase our evaluation of life, but this evaluation increases less and less as we grow richer. And we have a formula from Sacks, Stevenson and Wolfers (2010) for this which tells us that doubling incomes takes us about a quarter of a step up-wards on the ten-step Cantril ladder.[62] The relationship is now well-established across countries and seems increasingly supported by the evidence over time.

The main reason for posing this question is presumably to ask if the attention given to incomes and economic growth in assessing the human lot is well-founded. We are asking this now because we have far more data with which to answer the question and, thrillingly, have papers by Clark and Senik (2010), Deaton (2008), Stevenson and Wolfers (2008), and Sacks, Stephenson and Wolfers (2010) that do so.

Before going any further, it is worth making a distinction that all of the papers do as well (but never in their titles) – which is that what we are assessing is an evaluation of satisfaction with our lot. Happiness is argued to be rather different – a mood or affect, more fleeting and transitory. Happiness is measured as a daily experience – and seems to correlate less well with objective circumstances including income. It is clearly more volatile and not persistent. Daniel McMahon in *Happiness: A History* points out that the words for *happiness* in both ancient Greek (*eudemonia*) and every Indo-European language include, at the root, a cognate for "luck." In English, it is "hap", or "chance" – as in happenstance, haphazard, and perhaps. Etymologically speaking, happiness is being lucky.

Admittedly, it is not nearly as satisfactory a sound bite to ask if economic growth buys satisfaction or improves our evaluation of our lives – but that is where the strongest evidence lies. So the question properly phrased is: Does economic growth improve our evaluation of our lives? The answer from Sacks,

[61] See Table 1 (Sacks Stevenson and Wolfers, 2010) where the range of coefficients of life satisfaction on ln (household income) is reported as 0.21-0.28. These estimates pertain to the sample of developing countries.

[62] Life evaluation is assessed (for instance in the Gallup Poll) using Cantril's Self-Anchoring Scale (1965), as follows: *"Please imagine a ladder with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you, and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?"*

Stevenson and Wolfers (2010) is an unqualified Yes. Clark and Senik (2010) are more cautious but will not argue against it. Kahneman and Deaton (2008) offer the careful contrast between measurements of well-being obtained from Cantril's ladder of life, satisfaction with life (which asks how satisfied the person is with his or her life on a discrete scale) and emotional well-being or experienced happiness (the frequency and intensity of experiences of joy, anxiety, sadness and anger every day). They suggest that the first offers the closest approximation to overall assessments of well-being, while the second captures elements of both life evaluation and joys experienced yesterday. It is not surprising then that the correlates of emotional well-being or the immediate joys and pains are different from that of life evaluation. It seems clear that if we are to focus on a measure of subjective well-being, Cantril's ladder seems to be the most fruitful measure to pursue. It appears to be stable yet sensitive in response to small fluctuations and behaves in similar ways across varied contexts. So, as a measure, it seems to have properties that make it amenable to examination.

All of these papers take a hard look at the Easterlin paradox: that despite large increases in income *per capita*, average evaluation of well-being stays stagnant for a long time – or money does not buy well-being, measured on Cantril's ladder. The data now appear to say that i) within countries, the rich are perched higher on Cantril's ladder, ii) across countries, richer countries report higher average positions on the ladder and iii) the limited time series data says economic growth raises your position on the ladder too. The size of the effect is strikingly similar within and across countries and over time. The conclusion seems to tip us in favour of accepting that higher incomes make us more satisfied with our lives. I shall treat this as the maintained hypothesis and happily accept it on the available evidence.

The next real question, in vulgar parlance, is "So What?". So what if economic growth raises well-being thus measured? Does this mean that economists should now pursue the study of this new and possibly elusive goal – but policy makers should stop worrying and start being happy since incomes can be targeted without guilt?

The discussion here addresses three issues that pertain to the "So what?" question posed above.

What do we measure with Cantril's ladder? Is it utility?

What are we measuring when we obtain an evaluation of life? Does it measure our evaluation of all the conditions of our life, taking on board the constraints we face in constructing our best possible life? Or is it a relative measure, where the top rung just measures the best possible life given the constraints we face (such as discrimination, poor access to opportunities and so on?).

Is it utility? Does it offer the promise of a general measure of welfare? On this question, we have Becker (2008), Deaton (2008), Fleurbaey *et al.* (2009), and Sen, all saying Nay with Lord Layard firmly on the side of the Yea.[63] But it

[63] Layard argues that costs and benefits of policy interventions could perhaps be expressed in utils of happiness. He adds, "*But for the present the money equivalent of a util will do fine, provided it is specified as the extra money which would in the long run secure for the average person an extra util of happiness*". Sen calls this position "all Bentham and no Mill".

is relatively uncontroversial as an argument in the utility function: life satisfaction is something we value, and it is undoubtedly one of our objectives in life. The evidence tells us that money (although not money alone) can help us buy it. It is only one argument in the utility function – if utility reflects our overall objectives in life. In this vein, Becker argues that one could use the pursuit of (good) health as an analogy. Utility rises with expectations of living longer and being in better health. But we do not see people pursuing the goals of living longer and being in good health alone; we see them trading lower life expectancy for higher income by taking risky jobs and for the pleasures of indulgence. The same might be true for happiness. Perhaps it is more useful to conceive of subjective well-being as something that is often traded off against other goods that affect our overall utility or well-being?[64] It is clear that the appeal of subjective well-being lies in its ease of measurement. However, the fact that it is easy to measure does not mean that it is a satisfactory normative guide – it runs into the usual welfarist difficulty that aggregation will not respect individual preference orderings. It would also pose difficulties for a normative guide in a non-welfarist tradition for there is more to life than satisfaction with life.

Should subjective well-being (like the MDGs) be an object of targeting by policy makers?

Perhaps this question comes too late – the politicians have already answered this. The Office of National Statistics in Britain is to include questions about subjective well-being (though it has not yet revealed what these might be), the Sarkozy Commission has advised that the French do so and the Bhutanese already do so.[65] This measurement pits Voltaire against Dr. Johnson: whether all change must generate happiness as Dr. Pangloss claims *versus* Johnson's focus on whether humanity is capable of happiness at all. If we believe Dr. Pangloss, then measuring subjective well-being will tell us all we need to know – and if we understand Rasselas, we would despair of the likelihood of capturing anything with such measures.

The key issue is the need to distinguish the normative appeal from the positive use of this measure. The debate is about whether subjective well-being might be an argument in the social welfare function of society from a normative point of view – might we allow our social planners to use it as a guide to policy? Should we be willing to trade off nutrition for more satisfaction, for instance?

Clearly both Miss World and our politicians would like to raise happiness. Should we allow them to do so? The increased salience of the happiness and life satisfaction measures for

[64] Dercon, Krishnan and Krutikov examine rural to urban migration in India using data from the ICRISAT village studies. As with many studies of migration, they find that migrants rate themselves lower on Cantril's ladder (and on life satisfaction) than those left behind, despite higher consumption and incomes. They suggest that a model of risk-sharing across the extended family best explains this outcome: migrants are compensated with extra consumption for their lowered satisfaction with life.

[65] A cynic might point out that on well-being measures and the consumption of leisure, France might score more highly than on GDP alone.

politicians might well mean that just as reforms of the welfare state in Europe are accompanied by attempts to offer draconian incentives to unemployed single mothers to get back to work, we might soon have equally sharp incentives to direct us to activities that might make us happy and provide benefits for doing so. So should the grumpy middle-aged man in a suit be compensated for the horrors of his existence in the same fashion that we seek to protect the poor? It also offers the possibility of the not-so-benign dictator in the developing world using this agenda to suggest that his peasants may be poor but at least they are very happy. In sum, I am unpersuaded about whether we should offer our (less than) benign policymakers this extra lever to manipulate.

However, it is easier to accept the positive case for measuring and monitoring such measures perhaps. It would allow us a better grasp over both what the measures capture and the pecuniary and other variables that might affect it. It can be done quite simply and relatively inexpensively. The measure using Cantril's ladder (or satisfaction) appears to be relatively stable yet sensitive (see Atkinson *et al.*, 2005, Horley and Lavery (1991) and Howell, 2008).

Even the positive agenda must be accepted with care: we have evidence from Deaton showing that HIV prevalence has little effect on Africans' life satisfaction and suggesting that such measures miss crucial dimensions of well-being. Krishnan and Krutikov (2010) study non-cognitive skills of young people in the slums of Bombay – and find that the overall evaluation of life is influenced strongly by self-efficacy (a measure of control over one's life), which drowns out the effect of wealth on the overall evaluation of life. A similar result is to be found when examining the evolution of life satisfaction of Ethiopians: control over one's life is a strong influence on satisfaction with life. More striking, a quarter of all Ethiopians [66] surveyed who said they were satisfied or very satisfied with their lives would also be very unwilling to live their lives over again.

It is the case that we do not yet know what we might be capturing here and whether it is affected by incomes or whether the more satisfied are also prompted to create higher incomes and that brings me to the last question.

Is the relationship causal? Does economic growth raise well-being?

Does economic growth raise well-being – or does increased well-being raise incomes – or is there a mysterious third that conspires to raise both? It is clear that this is unresolved in this literature. There is undoubtedly scope for a growth industry in this area, with the usual pursuit of instruments and the attempts to randomly spread happiness that we can only applaud.

The agenda ought perhaps to be determined by our interest in a positive rather than normative point of view. More research into the various components of subjective well-being and their correlates are likely to be illuminating and interesting in their own right. As Clark and Senik suggest, it might offer the chance to investigate potential trade-offs and the marginal rates of substitution with various outcomes that accompany growth in incomes.

[66] This is obtained from the Ethiopian Rural Household Survey 2006.

Again, perhaps one ought to advise caution in developing a metric based on subjective well-being for we cannot reliably aggregate preferences in this dimension any more than in any other. We might also ask what if there was no connection whatsoever but, as Clark and Senik point out, we would still wish to keep economic growth on the agenda or as we say in English, *plus ça change*...

A summary

We have been offered a very clear account by Clark and Senik of the current state of play in research on subjective well-being. The evidence suggests that the measurement of (some) indices of subjective well-being may well be useful. These measures are relatively easy to collect – and in parallel with measures of pecuniary well-being, offer us the ability to assess well-being in a broader fashion than we usually do. However, there is clearly much more to investigate – which will make the researchers very happy – and much to learn, which ought to make policymakers cautious. Perhaps we might go further and urge that we be cautious of policymakers too: Tim Harford recounts the story of Sir John Cowperthwaite, the financial secretary of Hong Kong in the 1960s who refused to let British civil servants collect economic statistics for he feared it would merely encourage them to meddle in Hong Kong's success.

Conférences & Séminaires

References

Atkinson, T. (2005), "The Stability and Validity of Quality of Life Measures", Social Indicators Research Series, 26: 259-278.

Becker, G.S. (2008), "Comments on Stevenson and Wolfers", Brookings Papers on Economic Activity, Spring 2008: 88-102.

Cantril, H. (1965), *The Pattern of Human Concerns*, Rutgers University Press, New Brunswick, NJ.

Clark, A. and C. Senik (2010), "Does Economic Growth Raise Subjective Happiness in Developing Countries?", Paper for AFD-EUDN Conference, Paris 2010.

Deaton, A. (2008), "Income, Aging, Health and Well-being Around the World: Evidence from the Gallup World Poll", *Journal of Economic Perspectives*, 22: 53-72.

Dercon, S., P. Krishnan and S. Krutikov (2010), "Risk-sharing and Migration in Rural India", Queen Elizabeth House, University of Oxord, mimeo.

Ehrhardt, J.J., W.E. Saris and R. Veenhoven (2000), "Stability of Life-Satisfaction over Time", *Journal of Happiness Studies*, 1(2): 177-205.

Fleurbaey, M., E. Schokkaert and K. Decancq (2009), "What Good is Happiness?", CORE Discussion Papers 2009017, Université catholique de Louvain, Center for Operations Research and Econometrics.

Horley, J. and J.J. Lavery (1991), "The Stability and Sensitivity of Subjective Well-Being Measures", *Social Indicators Research Series*, 24(2): 113-122.

Howell, R.T. and C.J. Howell (2008), "The Relation of Economic Status to Subjective Well-Being in Developing Countries: A Meta-Analysis", *Psychological Bulletin*, 134: 536--560.

Kahneman, D. and A. Deaton (2010), "High Income Improves Evaluation of Life but Not Emotional Well-Being", *Proceedings of the National Academy of Sciences of the USA*, September 21, 2010, 107(3).

Krishnan, P. and S. Krutikov (2010), "Skill Formation in Bombay's Slums", Cambridge Working Papers in Economics No. 1010, University of Cambridge.

Layard, R. (2006), "Happiness and Public Policy: A Challenge to the Profession", *The Economic Journal*, 116(510): C24-C33, March.

McMahon, D.M. (2006), *Happiness: A History from the Greeks to the Present*, Allen Lane, London.

Mitra, S., A. Posarac and B. Vick (2011), "Disability and Poverty in Developing Countries: A Snapshot from the World Health Survey", *World Bank Social Policy Discussion Paper* 1109.

Sacks, D.W., B. Stevenson and J. Wolfers (2010), "Subjective Well-Being, Income, Economic Development and Growth", NBER Working Papers 16441, National Bureau of Economic Research.

Sen, A. K. (1999), *Development as Freedom*, Knopf, New York.

Stevenson, B. and J. Wolfers (2008), "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox", NBER Working Papers 14282, National Bureau of Economic Research.

Conférences & Séminaires

4. "Sustainable Growth": Do We Really Measure the Challenge?

4. Sustainable Growth: Do We Really Measure the Challenge?

Michel Aglietta, University Paris West (EconomiX), Cepii and Groupama – AM

Abstract

Macro-measurement is a long and arduous process that involves the system of national accounts. The challenge it raises is driven by the requirements of economic policy. Today, there is a need for a sustainable growth path that couples environmental concerns and development policy. This calls for nothing less than a sea change in national accounting, implying a shift from a system of income and expenditure accounts focused on GDP to a system of wealth accounts. The latter lays emphasis on an extended concept of capital, encompassing all assets that contribute to social well-being and an associated measure of "genuine" saving.

This paper analyses the problems involved in measuring different types of capital that are either ignored or barely dealt with per se by the accepted rules of national accounting. Yet, in developed countries, intangible assets are worth as much as productive fixed capital and constitute the most important factors of growth for the knowledge economy. Moreover, natural capital needs to be priced according to its scarcity given its function as a source of primary resources, an absorber of greenhouse gases and a means of conserving biodiversity.

Since the inherent nature of these types of capital means that they are measured at the discounted value of future rents, the choice of discount rates is critical to the valuing process. The discount rate is as important for estimating the cost of depleting non-renewable resources and the damage due to anthropogenic production of carbon dioxide as it is for valuing pension liabilities.

The paper shows that the process of valuing the different types of capital and estimating their substitutability in the production of social welfare is beset with radical uncertainty. There are multiple growth paths whose sustainability is open to question due to "unknown unknowns" in the interactions between economic and ecological factors, which may be linked to disruptive highly non-linear feedbacks. Therefore, the choice of discount rates is a deeply

ethical matter and needs to espouse a widely applied precautionary principle, given that the future of the generations to come is at stake. This principle holds true as long as there are situations of exposure to unlimited risk. Societies facing a catastrophic crisis of unknown probability need to engage immediate discussions about how to organise collective decisions on the appropriate policies to adopt.

Introduction: the milestones of progress in macro-measurement

From the 17th to the 19th centuries, there were many attempts to estimate national income and these efforts charted the way for macro-economic and demographic quantifications. They were, however, confined to individual enquiry that evaluated and compared the strength of nations and were largely inspired by mercantilist doctrine. In Great Britain, this type of research was pioneered by William Petty (1623-87) in 1665, followed by Gregory King (1648-1712). In France, the forerunner was Boisguilbert (1646-1714), who gave a very pessimistic account of Louis XIV's kingdom in his work *Le detail de la France* on account of the extremely unfair fiscal structure.

Yet, these scholarly studies never served as analytical tools for economic policymaking. In the liberal-minded 19th century, they attracted even less attention, to the point that from the mid-century the economic significance of macro-magnitudes was even cast into doubt. It was considered that economic adjustments in the business cycle stemmed from market interactions between individual agents and did not require state intervention. Even World War I made no impact on this conception of economics.

National accounting and macroeconomics are in fact the joint offspring of the Great Depression and World War II. The Great Depression had a crucial influence with its lasting mass unemployment, a phenomenon that was to spawn Keynes's General Theory. World War II provided the decisive impetus as there was a colossal growth in public receipts and expenditures and a pressing need to mobilise and re-allocate resources at the national level. There now emerged an overriding rationale for measuring intertwined economic aggregates in order to determine the overall level of economic activity.

Macro-measurement thus emerged as one of the tools for economic policymaking in the 1940s. Two seminal works published in 1940 encouraged its development for macro-management and for interpreting history: J.M. Keynes' *How to Pay for the War* and C. Clark's *The Conditions of Economic Progress*. A host of other authors followed, led by S. Kuznets, who pioneered quantitative economic history, along with J. Meade and R. Stone, who helped to standardise national accounting, jointly publishing, on Keynes' suggestion, a seminal article, *The Construction of Tables of National Income, Expenditure, Savings and Investment* (1941). The article presented their jointly elaborated structure of national accounting, which became the cornerstone of the 1945 memorandum for a System of National Accounts published by the United Nations in 1947. National accounting, Keynesian macroeconomics and regulation of aggregate demand are thus the three pillars of macro-measurement.

Adjacently but independently of the investigations on national accounting, the Harvard economist Wassily Leontief introduced input-output analysis in 1932. In 1941, he gathered the result of years of research into a book published by Harvard University Press: *The Structure of American Economy 1919-1929: an Empirical Application of Equilibrium Analysis*. Although his input-output table was not integrated into the main body of the Standard National Accounts, it was used by the United States Bureau of Labor Statistics, which designed the table for the year 1947. This work however came to a halt with the outbreak of the Korean War in 1950.

After 1950, the route was long to standardisation, international comparability and finally unification under the 1968 System of National Accounts, with a more complete system being introduced in 1993 under the auspices of the UN. The main problems to be tackled included the scope of economic activity for GDP estimates, the adjustment for price changes to capture volume magnitudes for GDP growth estimates, and the Purchasing Power Parity (PPP) converter to measure real GDP levels at international prices so as to achieve international comparability.

GDP was understood as being the crosscheck of three approaches to national income in what emerged as the development of flow accounting. On the demand side, GDP was defined as the sum of final expenditures. On the income side, it was the sum of wages, profits and rents. On the production side, it was the sum of the sectoral value-added. The standardised criteria designed by Richard Stone were pushed forward by Milton Gilbert at the Organisation for European Economic Co-operation (OEEC). He persuaded the statistical offices of member countries to adopt them as the basis for implementing the Marshall Plan.

However, the Anglo-Saxon view did not prevail for very long. The USSR had a much more restrictive notion of economic production based upon material output: its national accounts were anchored in material balances. Alongside this perspective, which reflects central planning in a society alien to a market economy, there was a French approach. Concerning the problem of valuing non-market services, French national accounts also used a more restrictive concept of production, although somewhat broader than the Soviet vision. The French system used input-output tables to obtain a more detailed analysis of the production structure, as well as providing a framework to describe financial operations. The system also described the financing behaviour of groups of economic agents and articulated operating, transfer and capital accounts. The compatibility between national aggregates and agent accounts was undertaken in the "*tableau économique d'ensemble*".

It was not before 1968 that the French and the Anglo-Saxon systems began to converge under the aegis of the UN. The French system enlarged its concept of production to encompass the production of public administrations. The Anglo-Saxon system opened up to agent accounts. There is a long-lasting debate on what method should be used to deflate nominal values. Competing indices can be used. The introduction of hedonic prices to allow for quality changes in commodities has never gained universal approval.

The problem of price changes over time mirrors the problem of spatial comparisons, which use the Purchasing Power Parities (PPP). The stumbling block lies in the definition of international prices. Measuring the prices of identical goods across a large array of countries requires an international standard bundle

of goods and comprehensive international surveys. By combining PPPs to eliminate cross-country differences in price levels and volume estimates over time, space-time comparisons were developed. The first studies on the subject were carried out by Milton Gilbert and Irving Kravis in 1954 and compared real expenditure levels in seven Western European countries with those in the United States. In 1968, Kravis, Heston and Summers launched a still ongoing international comparison project at the University of Pennsylvania. The UN Statistical Office has extended it and the OECD carries out such comparisons on a regular basis.

Meanwhile, the most ground-breaking use of time-space GDP comparisons is the monumental work of Angus Maddison at the OECD, who built time-series statistics of GDP in the world and its main regions stretching back to the year 1000, in order to assess the historical stages of economic development, the unequal distribution of world income, and periods of relative catch-up and decline.

Building on these earlier foundations, considerable headway has been made in tackling the many problems inherent to developing the system of national accounting, and in understanding the linkages between economic activity and social welfare, at least insofar as it is possible to apprehend social welfare in terms of economic value. Let us call it economic welfare.

4.1. National income and social welfare: a perennial problem

In his *Economics of Welfare*, Pigou explicitly formulated the question in 1920: can national income (NI) or national product (NP) be considered as an indicator of economic welfare? If so, economic values need to be interpreted within the framework of utility theory. An alternative solution favoured by Kuznets was to take into account the final objectives pursued by economic activity.

Each of these approaches comes up against some daunting obstacles. The utility approach attempts to use prices to integrate various measures provided by national accounting into welfare. However, the relationship between market prices and marginal utility is highly problematic, especially if one wants a cardinal measure of utility in order to aggregate individual preferences, notably because the comparison of interpersonal utilities is rife with insurmountable difficulties. Alfred Marshall, in his 1890 *Principles of Economics*, had already doubted that the utility value could be apprehended without recourse to market prices. The alternative approach does not carry much promise for an aggregate measure of economic welfare, since it has to clearly state the final objectives and find multiple "objective" standards of physical measure, whose integration into a single aggregate is either impossible or quite arbitrary.

Admitting that it is impossible to do away with prices, Pigou saw little point in hoping for an absolute measure of welfare using NI, but thought that some result could be achieved if ambitions were limited to *measuring the variation of NI* under restrictive assumptions. If the tastes of individuals and the distribution of their purchasing power within a group of people are fixed between times t and $t+1$, the economic welfare of such a group has increased between t and $t+1$ if the total income of the group has increased. This can be applied at the scale of the nation and measured by NI variation, as it can be assumed that that structural underlying preferences and income distribution do not change much in the short run.

However, if the focus is on social policies aimed at reducing inequalities and improving the well-being of lower-income citizens as a condition for better national social cohesiveness, a longer-term perspective is required. Not only can income distribution change, but also a narrowing in income inequalities might be considered as a highly effective means of improving social welfare.

Hicks entered the debate with a famous article published in *Economica* (1940) on "The Valuation of Social Income". Assuming that the budget of an individual with unchanged tastes, who consumes quantities (q) of different goods (i) valued at prices (p), such an individual is in a better situation in $t+1$ than in t if:

$\sum p_{t+1}q_{t+1} > \sum p_{t+1}q_t$. The individual could still afford to purchase the bundle qt but she prefers q_{t+1}. Conversely, the individual is in a better situation in t if: $\sum p_t q_t > \sum p_t q_{t+1}$. But the ranking of both situations is ambiguous if: $\sum p_t q_{t+1} < \sum p_{t+1} q_t$.

Hicks extends the income comparison to the whole of society. If the first two inequalities are satisfied together, the assumptions of unchanged tastes do not hold. Even if t+1 is a better situation than t, applying it to the whole society supposes acknowledgement that it is impossible for everyone to be placed at t in a situation as satisfactory as at t+1 by redistributing the quantities acquired during t.

Following Hicks's contribution, many attempts were made to find welfare indices that were beyond all ethical criteria in the possible reallocation of goods. However, in later research, Hicks admitted that such indices do not exist. Taking the opposite approach, Amartya Sen in his article Real National Income, published in the *Review of Economic Studies* in February 1976, used an ordinal framework. He showed that the marginal dollar possessed by a poorer person has a higher marginal value than the same dollar possessed by a richer individual. This therefore means that distribution of income must be taken into account, as well as the population structure and size, on the basis of judgments that explicitly vehicle ethical values.

Given this elusive theoretical foundation, what can national accounting do to measure economic social welfare? First, only a watered-down vision of welfare is possible, in which the satisfaction of needs is considered as being the final objectives of economic activity. The main measurement questions are: how should the public services that are not directly provided to individuals – the so-called "defensive" public expenditures – be dealt with? How should household expenditures that are induced by the constraints of urban life be treated? How can environmental costs and expenditures be taken into account?

The 2009 Report by the Commission on the Measurement of Economic Performance and Social Progress, chaired by Joseph Stiglitz, is the latest and most comprehensive contribution. It enters into the general debate on this topic and acknowledges a string of deliberate choices that have been made in the system of national accounts (SNA) since the 1970s.

In the 1993 SNA, it is clearly stated that GDP is not a welfare indicator. However, drawing on the long theoretical tradition dating back to Pigou and Hicks, the SNA observes that there is a relationship between variations in production or total consumption and variations in economic welfare. Which prices should be used to calculate the aggregates so that they express the relationship better? When prices change, the cost-of-living index must be equal to the amount by which a consumer's income will be modified such that he/she maintains the same economic situation as before the price change. This is the theoretical cost-of-living index. It can be demonstrated that, if consumer preferences are homothetic and their utility function quadratic, the Fischer index (a geometric average of the Laspeyres and Paasche indices) coincides with the theoretical index. However, if consumers have heterogeneous preferences, the volume of final consumption does not approximate economic welfare. Two obstacles remain: the aggregation of preferences and the impact of externalities. What then is the best solution to obtain a relevant aggregate even so?

In 1970, William Nordhaus and James Tobin wrote a seminal article, Is *Growth Obsolete?*, which triggered a number of other studies. This work was undertaken outside the framework of national accounting. It was not before the 1993 revision that some results were incorporated into the system. For the purpose of this paper, we focus on defensive public expenditures and environmental externalities.

4.1.1. Defensive public expenditures

Defensive public expenditures are not direct services to households and do not enter present-day total consumption as they produce flows of services that span long periods of time. They must thus be dealt with as public investments of three different kinds. Security expenditures (police, prisons and army) are investments in social capital. Expenditures to expand and improve the health system are investments in human capital. Environmental expenditures to mitigate and accommodate climate change, to find substitutes for the depleting fossil energy resources and to preserve living species diversity are investments in natural capital.

The impact of economic activity, which degrades the quality and the quantity of these assets, will be treated as standard capital depreciation and depletion of the capital stock, thus entailing a reduction in net national income.

However, some defensive expenditures imputed to households cannot be registered as investments. One example is the measurement of commuting expenditures related to the constraints of modern-day urban life. This expenditure produces services that use capital and labour, but it is not in itself an investment. It should instead be included in an extended definition of household production that is not paid for by firms in the present organisation of society and thus represents a transfer of added value from households to firms.

4.1.2. Environmental services

Insofar as economic agents consider natural capital as a free good, the social costs entailed by its depletion are not imputed to economic activity. This involves a diminution of economic welfare that is both uncontrolled and unreported. The losses due to the deterioration in the quality of environmental services due to economic activity – essentially pollution – thus need to be measured as they represent losses in the value of natural capital.

However, flows of environmental services, although part of economic welfare, are not included in final consumption given that their elasticity of substitution for economic goods, which would make it possible to aggregate them with market services, is not known. In this paper, they will be taken as a separate variable, included in the cardinal economic welfare function along with final consumption, the elasticity substitution being an unknown parameter.

4.1.3. From social welfare to economic sustainability

The way forward thus lies in extending the concept of capital to all assets that contribute to the maintenance and expansion of economic welfare for society as a whole over time. This is essentially a dynamic concept of *sustainability*.

In such an all-encompassing concept of capital, public services are non-rival and non-exclu-

sive public goods produced by tangible and intangible assets and which can be used by all economic agents. These collective assets are owned by society at large.

The advantages of broadening the concept of capital in order to study development are considerable, as this means that the relationships between income and asset values can be systematically analysed. For instance, to what extent does the capital invested in health influence labour productivity and real income growth, thus fuelling a sustained virtuous circle? Conversely, to what degree does pollution depreciate human capital by harming health? In the case of environmental degradation, a distinction must be made between the initial impact, which creates losses with unequal impacts on the capital account of economic agents, and the investments to repair these losses, which are current expenditures appearing in the flow accounts.

The change of perspective brought by the sustainability approach raises questions about the intellectual property of capital, which in turn leads to accounting reallocations. Items that had been traditionally treated as intermediary inputs must be reclassified as investments in capital. This is primarily the case for R&D expenditures, formerly considered as an intermediary input but lately reclassified in the 2008 system as an investment in intangible capital. It is also the case for mineral exploration, which produces new knowledge about the primary reserves of subsoil assets.

Sustainability is thus the new frontier of development. For decades, capitalist accumulation was based upon the assumption that the use of non-renewable resources was cheap, that the stock of non-used resources was free and that growth was unlimited. This crude model has become less and less relevant, and all the more so since the tremendous growth of services has highlighted the paramount role of intangible capital as a source of productivity. Meanwhile, the depletion of natural capital is continuing through oligopolistic rent appropriation and geopolitical manoeuvres for the sake of securing supplies. Added to this, the looming costs of climate change are so uncertain and threatening in worst-case scenarios that the stakes are nothing less than the future of humankind as we know it.

The concept of sustainable growth is being defined and developed with the ambition of being all-embracing so as to bring about a far-reaching reform of economic policy. Sustainable growth will lead to a revolution in economic thinking, accounting, government policy and the organisation of finance. Broadly speaking, sustainable growth integrates into growth trajectories the long-term protection of the environment and hence the welfare of future generations. A flow accounting framework geared to measuring and enhancing GDP growth must be supplemented by a stock-flow accounting system for measuring genuine capital (*i.e.* total wealth), which is the resource base for producing future social welfare and thus enhancing total wealth accumulation.

4.2. The challenge of sustainability: from income to wealth accounts

Early *ad hoc* attempts to measure national wealth were made in the 1950s. They were presented and discussed at the 1957 International Association for Research in Income and Wealth (IARIW) conference. The first long-run series of data on the US (1945-58) are the work of Raymond Goldsmith, published in 1962 and 1963. Sectoral accounts covered tangible and subsoil assets. Intangibles were limited to financial assets, the whole being calculated using the perpetual inventory method. However, these attempts were too narrowly scoped and the measurements too uncertain to consider integrating them into the standardised system of accounting.

In the mid-1960s, John Kendrick undertook a vast historical study, published in 1976, on total capital stocks in the US, including human capital and R&D. He included intangible investment in education, health and social mobility, and counted R&D separately. Kendrick's result was startling. He showed that, in 1929, the value of human capital equated the value of all other assets and that this was 15% higher in 1969. This means that standard national accounting and, even more so, private business accounting grossly underestimate the factors of growth.

Theoretical motives for improving the measurement of capital were kindled in the 1980s with the formal modelling of endogenous growth theory by Romer and others. With the work of Aghion and Howitt, the approach broadened to incorporate Schumpeter's innovative growth model. Yet, capital changes due to productivity-enhancing investments were only sparingly integrated. Software was included, but the integration of R&D expenditures met a strong resistance from national accountants. There were also problems for mineral and oil exploration expenditures, as these did not represent the value of mines and oil fields computed from the value of the rent after subtracting all costs, including the depreciation of exploration expenditure. More generally, the principle of valuing an asset by discounting expected future income streams met with reticence from accountants due to the elusiveness of expectations and the arbitrariness of the discount rate.

Little surprise, therefore, that these efforts encounter objections, pitfalls and obstacles. Wealth accounts involve the measurement of assets and liabilities. Measuring wealth, however, is a tricky enterprise as wealth-related items span considerable and very different periods of time. Two related problems appear: how can wealth be valued? How can stock accounts and flow accounts be articulated? As we know it, an economic measure is defined in terms of value or, in other words, as the social contribution of any activity using

available resources. However, if stock variables are taken into account, this implies mixing values that have different measurement statuses: realised value against monetary payment, imputed value, discounted (*i.e.* anticipated) value, replacement value (opportunity cost). Disparity in status means that integrating income (flow) accounts and wealth (stock) accounts cannot be undertaken without recourse to conventions and approximations.

Integrated wealth accounts only came under consideration in the 1993 system of national accounts. In his comprehensive history of national accounting, André Vanoli (2002) explains that, outside of flows pertaining to standardised national accounting operations, other flows must be described in two accumulation accounts: the "*other changes in asset volumes*" account, which records changes in the substance of wealth not due to production, primary income, capital transfers and variation in asset prices, and the "*asset re-evaluation*" account, which presents real capital gains or losses due to variation in specific prices relative to changes in the general price level.

If these estimates have been made, it is then possible to establish a link between savings and changes in the net value of wealth:

> Net saving = Δ (net real value of wealth) – net receipts from capital transfers – other net changes in asset volumes – real gains or + real losses in asset holdings

The World Bank is the institution that has ploughed deepest into measuring total wealth as a tool for assessing the sustainability of growth paths in advanced and developing economies. It has drawn upon a pioneer work by Pearce and Atkinson (1993), *Capital Theory and the Measurement of Sustainable Development: an Indicator of Weak Sustainability.* From its first 1995 report (Monitoring Environmental Progress) to its 2006 landmark report (*Where is the Wealth of Nations? Measuring Capital for the 21st Century*), the World Bank has explored methodology and conducted empirical investigations with a view to changing development policies. But today, it is time for national governments to take over.

Development policies can be linked to sustainability to the extent that they are interpreted as general portfolio management strategies. To implement such strategies, national government needs to redeploy its statistical apparatus in a collective effort to estimate changes in the total wealth of the nation in terms of size and composition. It is a major task that is as demanding as the one that laid the foundations of national accounting between 1940 and 1950. Only a dire threat requiring the mobilisation of national resources is able to spur a government into a collective effort for a policy shift that requires forging new policy tools. World War II was one such immediate threat for which the cooperation of democratic nations was vital.

The problem today relates to a world challenge on an entirely different time scale. Climate change, depletion of tropical forests and oncoming scarcity of water and fossil resources are long-term challenges that carry the threat of extreme events in the decades ahead or in the next century. The threat therefore involves future generations even more than the present ones. The stakes stretch way beyond the horizon of the issues addressed in the political debate of elective democracies. The reason for this is that the Earth's ecological system involves high inertia; this, however, may become vul-

nerable to uncontrollable diverging feedbacks if unknown thresholds are overstepped. As a result, humankind faces a world of unknown unknowns that nonetheless depend on its own behaviour over the next few decades.

Will governments embark on a policy of massive investment in radical innovation at the expense of present and near-term consumption, in the hope that fresh renewable energy investments will increase the substitutability of environmental and economic goods in the longer run? Or will they wait in the hope of obtaining more information on climate change developments and natural capital depletion at the risk of allowing the onslaught of irreversible damage? For a better appreciation of this dilemma, governments need to enhance their knowledge of the processes that lead to the sustainability or unsustainability of their economy, which they have the duty to steer and regulate within highly diverse scenarios.

The advanced economies have moreover been severely weakened by the ongoing repercussions of the financial crises. The immediate responses to these have massively transferred debt burdens and associated solvency risks from the private financial sectors to the public sectors. At the same time, population ageing has amassed contingent liabilities for the future, which have been aggravated in many countries by severe shortfalls in pension funding. And yet, if the environmental challenge is to be met, this will require a sustained effort in public investment to modify both the size and composition of total capital and thus reduce its exposure to the risk of future resource shocks and catastrophic climate events. This means that the sustainability of public debt needs to have credible short- to medium-run consolidation programmes that do not jeopardize the investments needed to ensure the future sustainable development of the whole economy. Public finance sustainability is the priority for the next decade in many advanced countries whose environmental choices are also crucial for overall economic sustainability in both developed and developing countries. As all policies dedicated to sustainability rely on the simulation of uncertain future growth paths, they all encounter the same measurement problems. Accordingly, this paper will attempt to highlight the stakes of public finance sustainability before moving to the more general question of sustainability in growth regimes.

Definition of total wealth and sustainability

Development depends on total wealth, which is to say produced, human, social and natural capital. Sustaining total wealth is the key to viable growth regimes. The different forms of capital are defined in the following way:

> Produced (tangible) capital = equipment + structures + urban land
>
> Intangible capital = human capital + institutional infrastructures + social capital + net foreign financial assets
>
> Natural capital = subsoil assets + timber resources + non-timber forest resources + protected areas + cropland + pastureland

The sum of the three components is the real national wealth. In the present state of our knowledge, this is far from being comprehensively measured. Only a massive statistical effort mobilising government resources and international coordination will bring about decisive progress on this count. Future growth prospects are linked to changes in real wealth, which are also termed adjusted net saving (*or*

genuine saving). As explained earlier, there is a weak link between changes in real wealth and the variation of future well-being or, in other words, of social welfare conditional to adjustments for defensive public expenditures, household production and environmental services. A strong link would require restrictive assumptions for the stability of individual preferences, income distribution and relative prices reflecting marginal utilities. Box 1 describes the formal derivation of the sustainability condition from the social welfare function under the proviso of strong sustainability. However, formally deriving the sustainability condition from welfare economics, even if this cannot lead directly to a measure, gives deeper insight into the meaning of the approximations that need to be made to construct a measurement methodology. Box 1 sums up the derivation of the sustainability condition from a cardinal social welfare function.

Box 1 *From social welfare to economic sustainability*

As explained in the discussion on the relationships between national accounting and welfare economics, environmental services are not perfectly substitutable with private consumption in a cardinal social welfare function. The limited and highly uncertain elasticity of substitution between private consumption and environmental services is modelled following the recent paper, *Ecological Intuition versus Economic Reason*, by Olivier Guéant, Roger Guesnerie and Jean-Michel Lasry (2010).

Definition of the welfare function

V is a social welfare function in continuous time, function of two variables:

$C(\tau)$, an aggregate of extended private consumption embodying the value of social services that are directly allocated to households and consumed over the same period, recalling that non-environmental defensive expenditures are counted as public investments in human or social capital.

$E(\tau)$, the environmental services that result from natural capital.

V is an inter-temporal cardinal function:

$$V(t) = \frac{1}{1-\eta} = \int_t^{\infty} U[C(\tau), E(\tau)]^{1-\eta} e^{-\delta(\tau - t)} d\tau$$

δ is the pure rate of time preference and η is the rate of relative risk aversion.

The imperfect substitutability between private consumption and environmental services described by a CES function whose limited elasticity is σ :

$$U[C(t), E(t)] = [C(t)^{\frac{\sigma-1}{\sigma}} + E(t)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}.$$

The sustainability criterion

The strong sustainability criterion is that V is not decreasing:

$dV/dt \geq 0$

•••

•••

The final goods and services generating utility, from either consumption or the environment, are produced by combinations of all of the above-defined types of capital on the basis of the best technology available. The different types of capital decrease with their use for production and increase through investment (productive and human capital), through natural repletion if renewable (forests) or irreversibly decrease due to extraction if non-renewable (subsoil fossil resources). The output generated by the combination of all types of capital comprises consumption, environmental services and services of investment in reproducible capital. The allocation of output depends on the adjusted rate of saving (genuine saving) and the public rule that makes environmental services available in specific quantities. It is assumed that the allocation mechanism is such that *V is not an explicit function of time*. It follows that the stocks of the different kinds of capital at t+1 are determined by the stocks at t and by the permanent allocation mechanism. This assumption is not benign since it abstracts from uncertainty as far as the allocation mechanism is concerned.

Under this assumption, one can proceed from period to period and determine in principle the entire future evolution of the stocks of different types of capital and the flows of consumption and environmental services. If there are n stocks of capital K_i at time t (i=1, ..., n), the values of the macro-variables in the economy are determined at all future times $\tau > t$. It follows that U is determined for $\tau \geq t$ and V(t) is also determined. This can be written as:

$V(t) = V[K_1(t), K_2(t), \ldots, K_n(t)]$.

The strong sustainability condition requires:

$$\frac{dV}{dt} = \sum_{i=1}^{i=n} \left(\frac{\partial v}{\partial k_{it}} \right) \left(\frac{dk_e}{de} \right) = \sum_{i=1}^{i=n} = p_e I_e$$

where p_i is the contribution of the i-th type of capital to intertemporal welfare, *i.e.* the shadow price of capital K_i, and I_i the net investment in this type of capital.

The strong sustainability condition means that, if capital is valued at its "fair price", the variation of social wealth at t is equal to the variation of intertemporal social welfare. *The criterion for sustainability is that real wealth is not decreasing, i.e.* that genuine saving is ≥ 0.

This condition is very general. It does not require that the welfare function have the analytical form chosen above to show how environmental services can be treated separately from household consumption. The absolute value of V does not need to be computed; what does have to be computed is the change in real wealth. However for strong sustainability to apply, the prices required to compute the elements of wealth are shadow prices. The latter are not observable, but are rather the prices that would prevail if all types of wealth were commodities traded in competitive markets under perfect foresight. Given that many types of capital are not traded commodities at all, they are partly calculated as *discounted rents, i.e. the price that must be paid for the scarcity of the resource*. It is on this point that measurement problems become more complicated.

•••

•••

Furthermore, in the above equation, shadow prices are measured in units of utility per unit of capital. This is not convenient for empirical use. One type of capital that has an observed market price can be taken as *numeraire*. Let us suppose it is i=1. This posited price is equal to 1. The prices of other types of capital, expressed in this numeraire, become price indexes. Let us call W the total value of wealth expressed in this price system:

$$W_t = \sum_{i=1}^{n} p_{it} K_{it} \text{ and } \frac{dW}{dt} = \sum_{i=1}^{n} p_{it} I_{it}$$

The condition of sustainability is $dW/dt \geq 0$. This can be used to measure weak sustainability where the range of the different types of capital is the most extensive that can be measured and where prices are the best possible approximations of shadow prices.

Technological progress and population growth

The condition of sustainability can be defined in a slightly different way. By dividing both terms of the equation by the value of the first type of capital, one obtains:

$$\left(dV/dt\right)\left(\frac{1}{p_1 K_1}\right) = \left(1/K_1\right)\left(dK_1/dt\right) + \left(\frac{p_2 K_2}{p_1 K_1}\right)\left(1/K_2\right)\left(dK_2/dt\right) + \ldots + \left(\frac{p_n K_n}{p_1 K_1}\right)\left(1/K_n\right)\left(dK_n/dt\right)$$

The economy is sustainable if the sum of the growth in the volume of the different types of capital, weighted by their elasticity of substitution to the one type chosen as numeraire, is non-negative.

Now let us suppose that there is Hicks-neutral technological progress. This can be interpreted as the "knowledge" growth rate taken as the numeraire. Its rate of growth is the growth rate (γ) in total factor productivity (TFP). With neutral technological progress, the elasticity of output to knowledge is 1. Therefore, the growth rate of real wealth becomes simply the sum of TFP growth and the growth rates of the other types of capital:

$$\frac{1}{W}\frac{dW}{dt} = \gamma + \sum_{i=2}^{n} \frac{1}{K_i}\frac{dK_i}{dt}$$

The growth in the volume of the different types of capital, including TFP growth is measured and the values obtained are added together.

The formula is valid if the population is constant. If the population is growing at rate g, the sustainability criterion must be applied to calculate the growth of real wealth per capita, with the caveat of constant population growth and of a wealth distribution that is independent of population change:

$$\frac{1}{W}\frac{dW}{dt} - g = \gamma + \sum_{1=2}^{n} \frac{1}{K_i}\frac{dK_i}{dt}.$$

Since the variation of total real net wealth or genuine wealth is the net investment of society, the sustainability condition is that wealth is not destroyed insofar as there is sufficient adjusted saving or genuine saving to match net investment. The sustainability condition thus becomes: *the development path of an economy is sustainable if, at every point in time, the adjusted social saving (or genuine savings) is non-negative.* Should it become negative, this means that society is destroying its wealth.

The definition of genuine saving is the following:

> Genuine saving = economic gross national saving – fixed productive capital depreciation + change in value of human capital + change in value of social capital – depletion of mineral and energy fossil resources – net reduction of forests – damages due to pollution in $\approx CO_2$

At this point, the theoretical framework for development policies dedicated to the long-term goal of growth sustainability has been set out. There are formidable obstacles to implementing such a drastic policy change considering current practices, which have led to unsustainable credit dynamics within the financial system, threats to public debt sustainability, distorted income distribution and mounting ecological perils. From the measurement standpoint, public finance management is today's priority, since this will be crucial for funding the investment necessary to the long-term accumulation of genuine wealth. Next, the measure of the types of capital that have been left aside in the national accounting systems is a precondition for designing adequate investment policies. This hinges on more than just statistics. The choice of the long-term discounting factor, crucial for measuring most types of capital, raises very difficult ethical questions.

4.3. Sustainability of public finances

The sustainability of public debt raises complex measurement problems. Should gross or net debts be factored in? If one considers that net debt is relevant, what assets are acceptable on the other side of the balance sheet? On the liability side, there are contingent liabilities that do not appear in either the annual fiscal accounts or the national accounts. Should they be included in the public debt? How can they be valued? The problem has been dramatised and has raised bitter controversies about the future sustainability of public pay-as-you-go pension systems. Assuming these problems are settled, what sustainability criterion would be appropriate? Is imposing an arbitrary limit (60% of GDP in the euro zone) a relevant and legitimate policy?

To address these questions, it is useful to be familiar with public debt accounting (Box 2).

Box 2 *Public debt accounting*

The fiscal yearly operating and income account is:

$H-T+i\,F_{-1}=\Delta D+\Delta M$,

where H is the amount of public expenditures, T receipts from taxes, D the stock of government bonds, ΔD the net annual flow and ΔM the monetisation of the government deficit. Thus the left-hand side is the deficit and the right-hand side is its financing.

This accounting equation can be expressed as a percentage of nominal GDP:

$h-\tau+(i-\pi-g)d_{-1}=\Delta d+\Delta m+(\pi+g)m_{-1}$.

The primary deficit in percentage of GDP is not dependent on the capital market:

$b=h-\tau-\Delta m-(\pi+g)m_{-1}$,

where $((\pi+g)m_{-1})$ is the seignorage.

One can define the *ex post* real interest rate adjusted to growth:

$\rho=i-\pi-g$.

The debt dynamic is described by the differential equation in discrete time:

$b+\rho d_{-1}=\Delta d$.

This equation is solved by iteration for a debt lasting n periods:

$$d_t=E_t\delta_{t,n}d_{t+n}-E_t\sum_{j=1}^{j=n}\delta_{t,j}b_{t+j}$$

The discount factor at n periods in the future is:

$\delta_{t,n}=\prod(1+\rho_{t+s})^{-1}$

•••

•••

It is to be noted that *the discount rate for the public debt is the difference between the real interest rate paid on average on the debt outstanding (a function of the debt structure) and the rate of growth of the economy*. The discounted value of the debt at t is thus equal to the discounted expected value of the debt at t+n minus the discounted value of the primary deficits between t and t+n.

The condition of sustainability is obtained by letting $n \rightarrow \infty$.

The public debt is sustainable if its discounted value $\rightarrow 0$ whenever the debt horizon tends to infinity. The condition is:

$\lim n \rightarrow \infty \, E_t \delta_{t+n} d_{t+n} = 0.$

This is the transversality condition, which means that the debt/GDP ratio must follow a stationary path for public finances to be sustainable. It does not have to converge to any maximal predetermined value (60% or any other number). Its value depends on the profile of future primary surpluses:

$d^* = -\lim_{n \rightarrow \infty} E_t \sum_{j=1}^{j=n} \delta_{t,j} b_{t+j}$

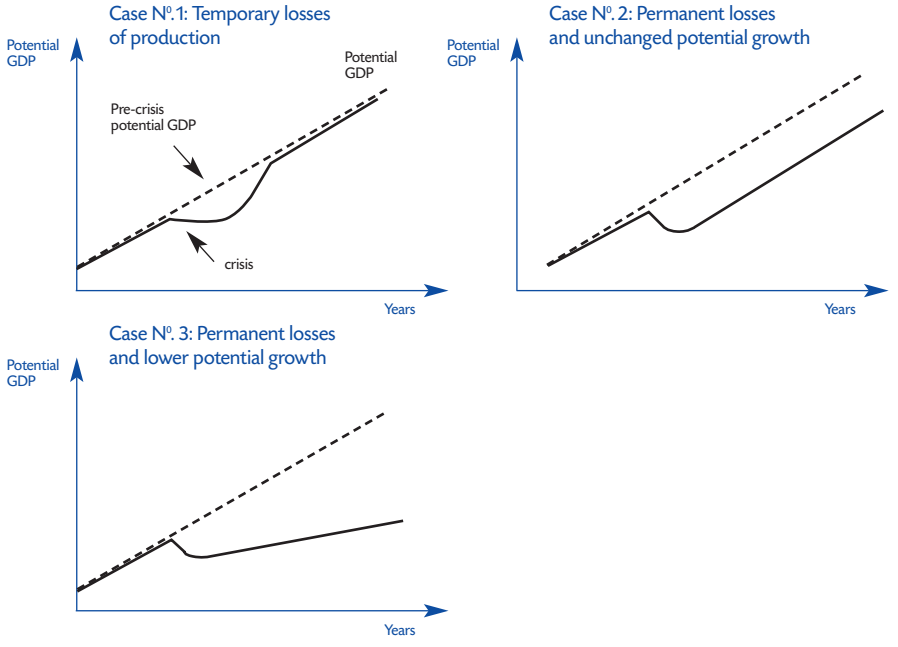
Box 2 shows that the present value of the gross sustainable debt in long-term equilibrium is the discounted sum of the expected future primary surpluses that finance it. The discount rate is all-important. The lower the discount factor (*i.e.* the higher the discount rate equal to the difference between the real interest rate paid on the public debt and the growth rate of the economy [r-g]), the less the future surpluses anticipated in the long run count in the value of the sustainable debt. In this case, the sustainability condition implies a near-term consolidation of public finances. Yet, a vicious circle may well set in. After a financial crisis, for instance, a large amount of debt is transferred from the private to the public sector. However, a fiscal effort that is too burdensome and too rapid could lower the trend growth rate of the economy for a considerable period of time, thus further lowering the discount factor such that the path of future debt becomes unsustainable and a default inevitable. The opposite arises if r-g is low. At the limit, under the golden rule (r=g), the discount rate is zero, meaning that all future primary surpluses are equivalent. The government disposes of an infinite period of time to straighten out its finances. Of course, if $r < g$, the public debt is sustainable even if the government continues to run up primary deficits, because future tax revenues are growing faster than the cost of servicing the debt.

The discount factor thus links the sustainability of the public debt to the larger condition of the sustainability of the growth regime defined above.

Figure 19 presents three scenarii that will have very different consequences for the future of public finances in the OECD countries most affected by the financial crisis. In order to design a credible medium-term programme for future primary balances, governments must infer from the history of financial crises

what the most likely future growth path will be. In turn, future growth will depend on the capacity of government policies to bring an endogenous influence to bear on growth.

Figure 19 *Scenarii of future potential growth*



To determine their medium-term consolidation plan, governments must have an objective d^*_{t+n} of their sustainable level of debt at the horizon (t+n) of their plan. This debt target must not be absurdly low. The European Commission's current proposal to return to the arbitrary ceiling of 60 percent of GDP, for instance, would surely be a self-defeating target that would destroy the credibility of governments. Whatever the case, once their targets have been fixed, the sustainability condition defined in Box 1 applies in the following way:

$$d_t - E_t \delta_{t+n} d^*_{t+n} = - E_t \sum_{j=1}^{J=n} \delta_{t,j} b_{t+j}$$

The left-hand side is the desired variation of the present value of the public debt. The right-hand side is the discounted flow of primary surpluses that need to be generated to finance it. As long as the equation remains valid, sustainability is respected at the pre-set target.

The fiscal effort required to satisfy the intertemporal budget constraint above is the difference between two ratios of fiscal revenues/GDP: the required fiscal pressure minus the present fiscal pressure ($\tau^* - \tau$). It is determined by the following equation:

$$\tau^* - \tau = [E_t \sum_{j=1}^{J=n} \delta_{t,j}]^{-1} [d_t - E_t d^*_{t+n} + E_t \sum_{j=1}^{J=n} \delta_{t,j} b_{t+j}].$$

This difference is such that the discounted value of the excess revenue due to the fiscal effort over the time period (t,t+n) is equal to the difference between the discounted value of the desired variation of the debt and the present value of future primary balances that would have been registered if past policy had been prolonged. As neither the tax system nor procurement expenditures are pliable to the government's will, the target level must not be set arbitrarily, but derive from a trade-off between the requirement of sustainability and the political feasibility of a change in policy. The longer the adjustment period, and thus the lower the discount rate, the more leeway the government has to implement a credible programme.

Whenever sustainability guides policy following a deteriorated budget and a large increase in the public debt, this means that a fiscal effort must be engineered since $\tau^* > \tau$. In order to determine realistic future primary surpluses, a complete balance sheet of the government sector is useful, as this is where future contingent liabilities and public sector assets must enter the picture (Table 8).

Table 8 *Government balance sheet*

Items	Assets	Liabilities
Fiscal assets and liabilities	Net present value of future tax revenues	Net present value of future primary expenditures
Fiscal assets and liabilities	Equity holdings in public sector companies Other financial assets	Gross public debt
Capital of the nation	Real assets	Net worth of the public sector
Public wealth	Total	Total

This presentation is a wealth accounting framework much like a corporate balance sheet. The whole nation is the community of "shareholders" of the public sector. The net worth of the public sector is thus akin to the implicit equity of the population, which represents the nation. The net worth is a measure of the nation's solvency and is broader than the measure of the gross debt/GDP ratio and in line with the problematic of sustainability. Whenever net worth is positive, the government can lower taxes, which raises the net wealth of households, bolsters investments and leads to increased public wealth and improved public services, with no detriment to the claims of government bondholders. If net worth is negative, public finances are not sustainable. If an outright default is declared, bondholders will absorb financial losses. If a consolidation policy is implemented to recover a positive net worth, taxpayers will suffer a

loss in their wealth through a higher tax burden or beneficiaries of public services will experience a welfare loss through downgraded public services.

In upcoming years, government net worth in advanced countries will be negatively affected by the lingering impact of the financial crisis plus the permanent loss of output if scenario 2, or *a fortiori* scenario 3 in Figure 19, comes to pass. Furthermore, ageing will gradually create contingent liabilities through a rise in health and pension expenditures. The cost of ageing on public finances depends heavily on long-term demographic and growth perspectives. These are linked to sustainable growth potential, the estimate of which in turn depends on the measure of genuine capital. If one admits that there are reasonably accurate procedures for measuring fixed productive capital using the perpetual inventory methodology, the most difficult task is to gauge intangible and natural capital.

4.4. Sustainable growth: measuring intangible capital

Measuring intangibles is a long-standing problem, even more so in business than in national accounting. In business, the concept of capital is highly restrictive. The so-called intrinsic goodwill of a firm is a kind of black hole, a source of profitability that is "guesstimated" and reported in financial markets in the event of a merger or acquisition. Besides, national accountants have been reluctant to capitalise expenditure on intangible assets. The accepted procedure is to treat expenditure on intangibles as intermediary expenses, not as investments that pertain to GDP.

However intangibles have been so boosted by the IT revolution that they have received much more attention over the last two decades. To recall the Solow paradox of the late 1980s: "you see the computer age everywhere but in the productivity statistics". The IT revolution developed the services industries, in which the measure of productivity depends on the ability of price indexes to reflect the qualitative changes of new products. Some countries tried to remedy this shortcoming by introducing so-called hedonic prices, but the conventions used to establish these are subject to debate. Things began to change in the mid-1990s. Yet, there remained a suspicion that the accounting treatment of intangibles, as intermediary products, underestimated labour productivity.

A broad estimate of intangibles for 1999 in the US, including software spending, R&D, human capital and the like, arrived at a stunning result. Investment in intangibles stood at around $1 trillion, roughly equivalent to the then investments in tangible capital (*cf.* Corrado, Hulten and Sichel, 2005). When brought to light, the intangibles hidden in the national accounts had a large impact on GDP, the rate of investment and labour productivity. This point-estimate raised various questions: do the biases introduced by ignoring such a large bulk of capital affect only the levels or also the rates of growth that they impact? If the latter hypothesis were true, how much growth is unaccounted for if intangible capital is omitted? What is the contribution of intangible capital to overall growth? What is the relative importance of capital accumulation and TFP when intangibles are taken into account? These questions are methodologically important since TFP is a residual value, the measure of our ignorance in understanding the growth process. Diminishing what is imputed to TFP is tantamount to an improved understanding of growth. Meanwhile, the system of national accounts has started to move forward, first by capitalising software, and then by accepting scientific R&D as capital expenditure. Yet this is only the tip of the iceberg. To understand the problems raised in valuing intangible capital, which meet with reservation from accountants, it needs to be explained why intangibles hold a special place in production and which categories of inputs define the domain under scrutiny.

4.4.1. Estimating intangibles.

Intangibles are characterised by their non-verifiable and non-visible nature, which makes it difficult or sometimes unfeasible to apply the perpetual inventory methodology. Some intangibles are non-rival or have non-appropriable returns. They are often developed in-house, offering no arms-length market transactions as a basis for their quantification. This is the case with brand promotion through advertising or other means (brand equity) or corporate management rules. Furthermore, separating price and quantity components is a tricky matter and even finding a suitable unit of measure is not self-evident. When there are no observable market prices, input costs are used as a way of circumventing the lack of direct observation. Another means of approximation is the use of a generic output deflator such as the non-farm business output price deflator should no specific price deflators be available. Approximation may be improved if labour is not the only factor used in producing the intangible. This lack of visibility stems from the absence of physical media as in the case of knowledge for instance. If knowledge is counted as capital, the link between investment expenditure and the capital stock is hard to identify since depreciation rates are elusive. However, this characteristic does not impact the way in which knowledge is used over time.

Non-rivalry raises the problem of measuring the marginal product of this type of capital. It could be zero in the direct production of the output – but improves production process efficiency and product quality. The non-appropriability of the return of some intellectual property means that its marginal product reflects only private benefits and costs. Nonetheless marginal accounting principles should apply to its valuation. It should be valued according to the present value of the discounted future expected income that it creates, the reason for this being what was described earlier with respect to welfare theory. *Investment is defined as any use of a resource that reduces current consumption in order to increase future consumption.* Consequently, all types of capital must be valued symmetrically, be it spending on R&D, employee training or plant and equipment. This means that the consumption side must prevail for unifying valuation principles whatever the differences in the production process and the practical difficulties of implementation.

How does the treatment of intangibles as intermediary products or as changes in capital modify national income? If spending on intangible capital is excluded from investment, it does not show up in the national income identity that measures the rental flow on capital by the residual value:

$p_k k = p_C C + p_I I - p_L L$.

Let us call N the volume of expenditure on intangible capital in a given period and p_N its price, R the volume of intangible capital and p_R its user cost. The national income identity is:

$p_k k + p_R R = p_C C + p_I I + p_N N - p_L L$.

The left-hand side of the identity is the non-labour payments accruing to both tangible and intangible capital. Overall profit accruing to both tangible and intangible capital is thus higher.

To estimate the volume of intangible capital of a defined category using the perpetual inventory methodology, it is necessary to calculate real investment by dividing the flow of nominal investment in a given period either by the average labour cost or by the non-farm business output deflator, depending on the nature of the category.

Human capital is embodied in people and, in this case, it can be assumed that the educational process earns the market interest rate for the length of time spent in education. Since the investment in human capital can be approximated by the number of years of schooling (abstracting from the intractable problem of accounting for differences in quality), human capital per worker is defined in steady state by: $H = he^{\rho A}$, where H is human capital per worker, ρ the appropriate rate of interest, h the real value of yearly education expenditures per worker and A the number of years of educational attainment.

The stock of human capital is:

total human capital = (human capital per worker)(number of workers), the number of workers being adjusted for mortality during working life. In competitive markets, the marginal productivity of capital equals its shadow price, which is equal to the real wage. One thus obtains: shadow price = (total real wage bill)/(stock of human capital).

For categories of intangibles that are closer to tangible capital, the perpetual inventory method can be used. Real investment N is estimated by dividing nominal investment by the non-farm business output deflator. Applying the capital accumulation identity, the stock of capital R is determined by:

$R(t) = N(t) + (1 - \delta_R) R(t-1)$.

How are δ and R(0) to be chosen? If δ is relatively high, experts can choose a benchmark year where R(0)=0. Rates of depreciation are estimated by field experts.

The user cost of capital still remains to be calculated according to the standard Hall and Jorgenson model for each category (i):

$p_{R,i} = [r + \delta_i - \pi_i] p_{N,i}$.

r is the competitive rate of return on capital, δ_i is the depreciation rate for asset I, π_i the expected capital gain or loss on asset i and $p_{N,i}$ is the investment price deflator.

4.4.2. Categorising intangibles

The corporate microeconomic approach and the developmental macroeconomic view produce widely differing lists of items to be grouped under the label of intangible investment. The corporate perspective is concerned with innovation, governance and social responsibility. At the broader macro-level applicable to development policy, the World Bank includes categories carrying a strong ideological dimension and promoted within the Washington Consensus. Some of these are not measurable but serve the purpose of promoting Western capitalism as the only optimal way forward.

Human capital is naturally common to both lists. Governance appears in both but with very different meanings. In corporate governance, what is at stake is the capacity to integrate business functions in line with one common goal and to control the managerial process. In the World Bank's view of governance, the benchmark is provided by US institutions. Institutional investments in other countries are ranked according to their distance from the US model. This preference lays emphasis on formal institutions. The rule of law is crowned with every virtue: social cohesion, political legitimacy and government effectiveness. This bias, however, completely ignores informal networks of social cooperation such as the Chinese *guanxi*.

Table 9 *Intangible investment and capital by category*

Business intangible investment	Development-related intangible investment
Computerised information	Human capital
Innovative property: - R&D based on scientific knowledge - Non-scientific commercial R&D	Social capital: - degree of trust in society - ability to work for common purposes
Investment in organizational capabilities: - strategic planning - redesigning existing products - investment in brand names	Governance: - judicial system - property rights - legitimacy of governments
Investment in firm-specific human competencies (management and professional)	Other assets: NFA + omissions in evaluating other forms of capital

4.4.3. Impact of business intangible capital on growth

In this section, we follow the work of Corrado, Hulten and Sichel (*Intangible Capital and Economic Growth, 2006*) relating to the United States, which is one of the most advanced pieces of research to date. They take GDP levels and growth rates in national accounts that treat most intangibles as intermediary products and compare these with the figures that might have been reached had the accounts posited the alternative assumption of intangibles as productive assets. Their methodology is described in Box 3.

This growth accounting exercise delivers impressive results. Capitalising intangibles increases the rate of growth of output per hour, both in the 1973-95 and 1995-2003 periods, such that the level is 20 percent higher than in national accounts at the end of the first period and 11 percent higher at the end of the second period.

The role of capital in labour productivity is substantially increased when intangibles are included. Conversely, TFP growth declines in importance. The Solow paradox is resolved since the role of capital in the acceleration of productivity growth linked to the IT revolution is much larger with intangible capital.

Last but not least – and contrary to well-anchored beliefs – firm-specific resources (organisational competencies, firm-specific human competencies and commercial innovative property) have had the most impact, while scientific R&D has had a much lesser impact.

Box 3 *Intangible capital and growth accounting*

The notations have already been defined earlier. Two approaches to growth accounting are compared: one that classifies intangibles as intermediate input and the other as capital.

Intangibles as intermediate input

Three branches of the economy produce goods C, I and N with two factors of production, L and K, allocated to all three goods and N to C and I.

The production functions are:

$N = F_N(L_N, K_N)$
$I = F_I(L_I, K_I, N_I)$
$C = F_C(L_C, K_C, N_C)$.

Assuming that the factors of production are paid the value of their marginal productivity, the income identities are:

$P_N N = P_L L_N + p_k K_N)$
$P_I I = P_L L_I + p_k K_I + p_N N_I$
$P_C C = P_L L_C + p_k K_C + p_N N_C$.

Under the assumption that all intermediary input markets clear, N is both an output and an input for producing other commodities. It disappears in the aggregate of GDP identity:

$P_y Y = P_C C + P_I I = P_L L + p_k K$.

In conventional national accounting, the GDP growth rate (g_Y) is the sum of the TFP growth rate (γ) and the weighted sums of the contributions of the growth rates of the factors of production. The weights are their shares in GDP assumed to be equal to the corresponding output elasticities.

$g_Y = s_C g_C + s_I g_I = s_L g_L + s_K g_K + \gamma$.

Intangibles as capital

The output of intangibles appears in the production functions of C, I, N as cumulative stock R.

The production functions are:

$N = F_N(L_N, K_N, R_N)$
$I = F_I(L_I, K_I, R_I)$
$C = F_C(L_C, K_C, R_C)$.

The income identities in the three branches are:

$P_N N = P_L L_N + p_k K_N + p_R R_N)$
$P_I I = P_L L_I + p_k K_I + p_R R_I$
$P_C C = P_L L_C + p_k K_C + p_R R_C$.

The GDP identity:

$P_y Y = P_C C + P_I I + P_N N = P_L L + p_k K + p_R R$.

•••

•••

GDP is expanded to include the production of the newly produced intangibles on the production side and the flow of services from the stock of intangibles on the income side. P_R is the user cost associated with the services of the intangible stock. This is a source of income that is absent from the intermediate input treatment. GDP is thus more comprehensive and larger in magnitude. GDP being larger, the shares of the factors of production are recalculated.

$g_Y = s'_C g_C + s'g_I = s'_N g_N = s'_L g_L + s'_K g_K + s'_R g_R + \gamma'$.

TFP growth and labour share are lower than they were following the former accounting rules. They are reduced by the coefficient

$$\lambda = \frac{p_C C + P_I I}{p_C C + P_I I + p_N N}$$

The capital share is larger because it includes the share of income that accrues to intangible capital; hence, the wage income in human capital is counted as capital income.

4.5. Sustainable growth: measuring natural capital

Natural capital is extraordinarily diverse. The World Bank distinguishes subsoil assets, forests and land resources. Subsoil assets are non-renewable fossil deposits: coal, oil, natural gas and minerals. Land comprises cropland, pastureland and protected areas. Forests encompass timber and non-timber resources. Other natural resources, like air and water, cannot be assigned a monetary value. Attempts have been made to register them in energy materials inventories: wild flora and fauna, wild fish stocks in oceans, continental waters, etc.

The main problems for accounting purposes involve the depletion of natural resources and the degradation of non-market natural wealth. As for the latter, monetary evaluation in satellite accounts is carried out for damages caused by economic activity to biodiversity. Marketable natural resources are either non-renewable (energy and minerals) or renewable (forests and fish).

Prior to 1993, non-renewable subsoil resources were recorded in production accounts based on the date of extraction. On the contrary, newly discovered resources were not recorded at the date of discovery. In the 1993 system of national accounts, discoveries and takings on non-produced assets are recorded in the other changes in volume account.

All non-renewable resources generate rents, which provide the basis for their valuation. In the case of renewable resources, the stock is infinite and the intrinsic value of the harvested resource is zero if the quantity of stock taken is less than the natural growth. In this case, the total sales value of the quantities harvested are imputed to the production of forestry or fishing. In the opposite case, when depletion of the stock exceeds the natural growth, the exploitation of the resource is not sustainable.

Therefore, net investment in forests is equal to: natural growth + (replanting – amount depleted).

A rent appears if: amount depleted > natural growth.

The resource acquires a scarcity value or, in other terms, a positive monetary rent. It should be treated in the same way as the rent of non-renewable resources (*cf.* Vanoli, 1995). Fish stocks are excluded from this analysis because overfishing may be due to that demand temporarily exceeds fishing capacity. It is not the intrinsic rent but the costs that increase as long as the resource remains depleted.

All in all, the scarcity value of a renewable resource can be measured by the economic cost of bringing the asset back to the level immediately before it was overexploited. This cost is the discounted value of the loss of income due to the limitation or absence of exploitation while the resource is being replenished.

The framework used by this paper to model sustainable growth as relating to social welfare requires that natural capital be measured since the quality and quantity of the environmental services this capital provides diminish with economic activity. *Depletion in the stock of non-renewable resources, overexploitation of renewable resources and pollution generate losses in the value of natural assets which reduces genuine saving.* As a result, they impact the variation of intertemporal social welfare.

4.5.1. Measuring the value of non-renewable assets

The value of subsoil assets has to be estimated indirectly since there is no transaction that takes into account the total amount of oil or mineral reserves in oilfields and mines. For this estimation, the economic resource rent, which is the net receipt from the exploitation of the underground resource for the total period of extraction, must be calculated. The value flows at the present date must then be discounted.

The rent is the difference between the market value of the primary resource after extraction and the total of the costs incurred for prospection, development and extraction, including the normal return of the productive capital invested in the exploitation.

The rent is a pure scarcity value. Contrary to a produced input, the rent of a non-produced input does not remunerate any factor of production, since no factor has produced fossil resources. It is the value of a gift of nature created by the human activity that uses it. In its 1997 report, *Expanding the Measure of Wealth: indicators of Environmentally Sustainable Development*, the World Bank defines the economic rent of a natural capital as *"the inherent surplus value in the extraction or harvest of a resource"*. This is why it is said to be the inherent value of the resource regardless of the owner, on condition that a clear distinction is made between the owner (public or private), the prospector and the extractor. This "absolute" rent should not be confused with the differential rents arising from productivity differences in the costs of the different extraction processes and locations.

In national accounting the rent is measured as follows (*Accounts for Subsoil Assets*, Eurostat 2000):

> Rent = extraction value at base price – intermediary input – wage compensation – net taxes on production – fixed capital depreciation – normal return on fixed capital (including intangible asset for prospection expenditures).

The taxes to be imputed are taxes non-specific to the industry. Specific taxes are integrated into the rent. The main problem in estimating the rent stems from the net stock of fixed capital in the extractive industry and the choice of a "normal" rate of return for this stock. When the total rent is calculated, the share accruing to the state is determined, with the share appropriated by the extracting firms being obtained by difference.

The value of the stock (V) is the present value of the discounted yearly rents (R) over the whole period of extraction:

$$V_t = \sum_{\tau=t}^{t=n} \frac{R_\tau}{(1+r)^{\tau-t+1}}$$

The crucial assumption is the discount rate. The World Bank has chosen a controversial social rate of return: $r = \delta + \eta g$, where δ is the pure rate of time preference, η the elasticity of

the marginal utility of income and g the potential rate of growth of consumption *per capita*. Eurostat is even more open to criticism for its choice of a rate close to the average yield of sovereign bonds in advanced economies. More will be said on the discount rate in the last section of the paper.

There is a simplistic and straightforward way of cancelling out the influence of the discount rate. This is the Hotelling rule, which assumes that the relative price of the scarce resource increases at a rate equal to the discount rate. In that case, the value of the stock is simply: $V=\rho Q$, where Q is the quantity that has been extracted and p the unit rent.

More generally, the rent has the perverse effect of transferring value to non-producers. The capital gains of the rentiers are:

Capital gains = (rate of increase in real rent) (resource stock)

Total capital gains = $\sum$ capital gains over resource-owning countries = world capital losses to consumers

Rent has always been a curse for the growth of capitalism. The increase in the real value of rents over time raises the real price of the resource for users. Hence, their real wealth declines, which gives rise to a negative wealth effect on consumption. It is also a curse for the countries that own the primary resources. Most countries with a high share of resource rents, such as Nigeria, Zambia or Venezuela, have experienced a long-run decline in real *per capita* income. Asian countries with scant depletable resources, like Korea, Thailand or India, have enjoyed high rates of capital accumulation and income growth.

There are three possible negative impacts of high levels of rents. The first is Dutch Disease, in which the overvaluation of the currency thwarts industrial profitability and results in low investment rates and lagging labour productivity. The second is the fact that high rents stall reform initiatives. The existence of a powerful rentier elite hampers the emergence of an entrepreneur class and dissipates the rental value, diverting it away from investment in education, health and infrastructure. The third is resource price volatility, which discourages investment in exploration and in improvement of extraction and processing technologies. Not all countries fall into this trap. The government of Norway is pursuing a focused and generalised policy of reinvesting the rent, so as to substitute intangible capital (financial and knowledge-intensive industries) for natural capital (oilfields).

Agricultural land is measured in the same way as subsoil resources:

Value of cropland = present discounted value of land rents

Value of pastureland = opportunity cost of preserving land for grazing

Urban land is not treated in the same way and is conventionally computed as percentage of total productive capital.

4.5.2. Measuring damage to environmental capital

All the different categories of damage should be measured and deducted from wealth accumulation in order to arrive at a satisfactory measure of genuine saving. However, given the present blockade-type situation created by powerful business interests and the unwillingness of governments to acknowledge the gravity of the risks, damage through natural resource depletion and air and water pollution are grossly underpriced.

The least recognised form of resource depletion is the destruction of the tropical rainforests, which, along with the oceans, constitute the planet's most important carbon sinks. However, as a recent report of *The Economist* (Seeing the Wood, September 25, 2010) noted, "*the importance of plants' ability to store carbon in making the planet habitable is still not widely appreciated*". About half the Earth's original forest area has been cleared and the pace of deforestation has accelerated in the last 60 years. Rainforests play a fundamental role in the sustainability of ecosystems. They capture and store carbon dioxide, and do so increasingly when the atmosphere is carbon-heavy – a process known as the carbon fertilization effect. They are also home to more than half of the animal and plant species on earth, and the source of a large variety of staple foods and medicinal plants. They regulate water run-off, mitigating the risk of flooding and droughts, and while trees increase rainfall, deforestation reduces it.

Forest degradation is still worsening in the tropical regions (Amazonia, Indonesia, Congo) and in the boreal taiga. In these regions, forests are threatened by global warming and the irresponsibility of governments. Global warming causes the permafrost to melt and releases billions of tonnes of methane. It also gives rise to calamities that are particularly dangerous for the integrity of forests (aridity, droughts, pests and fires) by causing deforestation, which in turn aggravates such phenomena through a self-sustaining feedback mechanism. Such vicious circles can bring the rainforests close to the so-called "tipping points" where they become ecologically unviable.

Furthermore, governments do not have demographic policies, which means that in forty years' time the world's population is predicted to rise from 6 to 9 billion. Given the currently low level of agricultural productivity in Africa, the rising demand for biofuels, palm oil and soybeans, along with the fast-growing demand for food accentuated by the middle classes' shift to a meat diet in emerging countries, the pressure to free up both cropland and pastureland through deforestation will become even more acute.

To meet this challenge, the UN has launched the far too narrowly scoped Reduced Emissions from Deforestation and Forest Degradation (REDD programme). The idea is sound enough in itself: rich countries should pay poorer ones not to fell trees. They should do so because they need the rainforests in order to control their own climate. REDD's current budget, however, is much too low (USD4.5 billion) to have a significant impact on the tens of millions of farms that are settling in or close to forests.

To effectively put a term to the destruction of forests – let alone repair the irreversible damage that has already been done–drastic changes must be made to national policies: much better forest management, land reform, a sustained effort to improve agricultural productivity and much tighter law enforcement. None of these shifts can come about as long as forests are not properly valued.

In national accounts, tree clearance is attributed a value, whereas standing trees have no value! Accounting that reports clearance positively and ignores the multiple costs from externalities that are the side effects of clearance makes no sense. Clean air and non-acid rains are public goods that everyone wants but which no one is prepared to pay for. It is the role of governments to set a price on the

non-felling of trees. *This non-use value should be defined as the opportunity cost of cutting them down and selling them.* In principle, the opportunity cost should represent all the environmental services provided by the rainforests, from water regulation to carbon sequestration and biodiversity protection.

One way to value, be it only partially, this opportunity cost with a view to halting deforestation is by pricing at least one of the main negative externalities, which is the release of carbon dioxide. Let us suppose that a farmer wants to clear a hectare of forest for pasture. One can calculate the amount of carbon dioxide that is released into the atmosphere as a result of the destruction of this carbon sink. For a given price of carbon, one can determine the cost to be assumed by the farmer, had he to pay for the marginal increase in emissions. If the cost is higher than the prospective gain from the projected farmland use, the tropical forest would be protected. However this mechanism is not feasible in every situation. In Africa, deforestation is often a matter of survival for the very poor, who need to cut down trees to ensure their heating and food. Carbon credits can indirectly channel financial resources to such local populations by funding micro-development projects. However, other instruments are also required for designing and implementing micro-projects (*cf.* de Perthuis, 2010).

Carbon credits are financial instruments for redistributing the costs of global damage from climate change. Indeed, all emissions contribute to global warming, but the impact of the average increase of global temperature is latitude-dependent. There is thus a single social cost of carbon. Financial mechanisms must be agreed upon internationally to finance carbon reduction where it is most needed.

Because CO_2 emission is a global externality that contributes to global warming whatever the geographic location of emission, the marginal social cost of carbon is the weighted sum of future marginal damages generated by the emission of one additional unit CO_2 equivalent at date t. The weighted coefficients depend on the discount rate and on the rate of natural CO_2 absorption by carbon sinks. The marginal cost of carbon is a value that rises with increases in the amount of emissions. The marginal cost of emission reduction (marginal abatement cost) is the cost incurred in order to reduce the volume of emission by one additional unit of greenhouse gases. It is a decreasing function of the amount of emissions. The cost-benefit analysis compares the marginal damage and the marginal cost of abatement at each level of carbon reduction. At equilibrium: *marginal damage = marginal abatement cost.* The price given by the equilibrium point is the *social cost of carbon.* This price should serve as the basis for pricing a carbon tax and for calculating the returns to innovations geared to reducing GHG emissions and to adapting housing and transport to climate change.

4.5.3. The social value of carbon

The problem is the practical quantification of the marginal damage function and the emission abatement cost function, as the undertaking is rife with uncertainties that limit the validity of cost-benefit analysis in integrated evaluation models. To account for uncertainty with respect to damages, the Stern review simulated numerous scenarios. Another method is to apply the precautionary principle to define an option value that makes allowance for the irreversibility of climate change. With time, new information will come to light and

should be used to structure a sequential decision process. A crucial factor for calculating the social cost of carbon emission, building alternative scenarios and developing a decision process that integrates the option value of irreversibility is the discount rate.

The idea is that the social value of carbon is universal (*cf.* Quinet, 2008). As such, its price should ideally be set at a climate conference under the umbrella of the UN. It is important that the full signification of this price be understood: it is a notional price that monetarises a public good, *i.e.* the improvement of an environmental service translated as each additional tonne of avoided CO_2 emissions. This price should thus be factored in when calculating the returns of long-term investment projects in order to highlight their social returns. But how can a common convention be reached when the attitudes to climate change are so diverse and no official international agreement is conceivable in the near future? How can this be decided? How can a common benchmark be established for financing investment in CO_2 emission reduction?

Most of the theoretical findings are not actionable as the models used are constructed under certainty conditions and technical progress is assumed to be exogenous. They nonetheless provide a basis for understanding the variables that play a key role and for determining the constraints to optimising the social utility function. Starting from this theoretical basis, two types of complexity need to be introduced: the shift to uncertainty and endogenous technical progress that takes innovation into account. This uncertainty is daunting since optimal control implies taking decisions that involve two conflicting irreversibilities. Environmental irreversibility results from unknown threshold and feedback effects that are potentially catastrophic for the ecological system should GHG concentration exceed such thresholds. This thus leads to recommendations for early intervention given the convex cost of damages. The irreversibility of technical progress, on the contrary, places an option value on waiting for more highly performing technologies to become available. It thus encourages a low level of initial effort.

Models using certainty conditions and exogenous technical progress do not raise such problems. They fall into two approaches depending on whether future damages can or cannot be calculated. The first approach is known as the *cost-benefit* or *Pigouvian* approach, in which the social value of carbon results from the equalisation of the marginal cost of reducing emission by one additional tonne of CO_2 and the discounted value of the future marginal damages resulting from the introduction of this one additional tonne into the atmosphere. This approach enables the social value of carbon to be calculated at any time on a path that depends on the specification of the functions of changes in GHG concentration levels, the impact of this concentration on the rise of global average temperature and the effects of this rise on economic and social damages. The second approach is dubbed *cost-effective*. This does not attempt to calculate the damages and introduces a constraint on maximum concentration not to be exceeded at a specified future date. It aims to find the means of achieving this objective with the least cost to society. This approach determines the pace of change of the social value of carbon, but it has the shortcoming of leaving the initial value undefined. The main results are shown in Box 4.

Box 4 *The social value of carbon under certainty conditions*

Cost-benefit approach

The function of atmospheric GHG concentration depends on the cumulative and non-sequestered stock from previous emissions.

Let Xt be the stock at time t, x_{t-s} the emission at the time t-s and a the rate of natural carbon absorption in the carbon cycle.

The stock level at t is: $Xt = \int_0^{\infty} e^{-as} x_{t-s}\, ds$

Let $D(X_t)$ be the function of damages resulting from GHG concentration and δ the social discount rate. The optimal social value of carbon (base price of CO_2) must be equal to the marginal cost of damage:

$V_t^{co_2} = \int_0^{\infty} e^{-(\delta+a)s} D'(X_{t+s})ds.$

Knowing the damage function is crucial. Stern takes the global damage function to be convex:

$D'(X) = \varepsilon X^{\gamma}$, where $\gamma > 1$ denotes the convexity of the damages. Assuming that emissions rise exponentially at rate g, the social value of carbon at any point in time can be calculated:

$V_t^{co_2} = \varepsilon \int_0^{\infty} e^{(\gamma g - \delta + a))S} ds = \frac{\varepsilon X_0^{\gamma}}{\delta + a - \gamma g}\, e^{\gamma g t}.$

Due to the convexity of the damages, this value increases at rate $\gamma g > g$, which is the emission concentration growth rate. Its initial level, moreover, decreases as the discount rate and natural absorption rate rise.

Cost-effectiveness approach

The reduction of CO_2 emissions is linked to the depletion of fossil fuels. Let Z be the concentration cap with a target of reaching this limit at time T1 (*e.g.* 2050) without ever exceeding it. There are thus three time periods. From 0 (=1990, depending on the basis used to calculate concentration) to T_1, concentration rises. After T1, concentration remains at its maximum limit, which allows the carbon value to gradually decrease while fossil fuel stocks are being progressively depleted. At T_2 (≈2100), the carbon value has fallen back to 0 and it is the scarcity rent that increases, which generates returns to substitute low-carbon energy sources (assumed to be available with no prior innovation cost) until time T_3 (≈2200), when these clean energies will be the only ones used.

We now calculate the evolution of the social value of carbon over the period $[0,T_1]$, which directly concerns us. Let C_t denote fossil fuel consumption and U(Ct) the utility obtained from this. The control programme for social optimisation is:

Max $\int_0^{\infty} e^{-\delta t} U(C_t)_{dt}$ under the constraints on:

the depletion of fossil fuel stock S: $\dot{S}_t = -Ct$

the CO_2 concentration: $\dot{X}t = \varepsilon C_t - aXt.$

the concentration cap objective: $X_t \leq Z.$

•••

•••

This optimising control programme can be decentralised if markets are perfectly competitive, through dual prices that the public authorities can control by using taxes and subsidies as incentives for private actors to behave in line with the social optimum. These prices are the optimal dual variables or the Lagrangian multipliers of the generalised function that incorporates the constraints. Let λt be the scarcity rent (implicit price of the fossil resource), μt the carbon value (implicit price of the atmospheric carbon stock) and ω_t the multiplier associated with the point at which the concentration constraint is reached.

The values of the dual prices evolve according to the differential equations:

$$U'(C_t) = \lambda_t + \varepsilon \mu_t$$

$$\frac{\dot{\mu}_t}{\mu_t} = \delta + a - \frac{\omega_t}{\mu_t}$$

$$\frac{\dot{\lambda}_t}{\lambda_t} = \delta$$

As long as the carbon concentration constraint is not saturated ($t < T1$), the scarcity rent for fossil fuels increases in line with the discount rate (Hotelling's rule) and the carbon value in line with the sum of the discount rate and the natural absorption rate ($\omega = 0$).

Introduction of exogenous technical progress:

Let Y(t) be the per capita consumption of private goods and services and $\Gamma(A(t),t)$ the cost of *per capita* abatements. The optimising policy is written:

$\text{Max}_{/A} \int_0^{T_1} e^{-\delta t} U(Y_t - \Gamma(A(t),t))dt$ under the constraint of not exceeding the concentration cap:

$$X_0 + \int_0^{T_1} (X_t - A_t)dt \leq Z$$

Designating Γ_{mt} as the marginal abatement cost and ω_t as the Lagrangian multiplier associated with the constraint, the first order condition is:

$$e^{-\delta t} U'(Y_t - \Gamma_t)\Gamma_{mt} = \varpi_t.$$

Thus, differentiating this condition with respect to time t we obtain:

$$\frac{1}{U'_t}\frac{\partial U'_t}{\partial t} + \frac{1}{\Gamma_{mt}}\frac{\partial \Gamma_{mt}}{\partial t} = \delta \text{ since } \mu = 0 \text{ for } t < T1.$$

The derivative of marginal utility compared to itself is equal to the opposite of the product of risk aversion (η) and the consumption growth rate *per capita* (private consumption – abatement cost) g. The marginal abatement cost, which is the social value of carbon, thus increases at a rate that is the economic discount rate of net capital return:

$$\frac{\dot{\Gamma}_m}{\Gamma_m} = \delta + \eta g = r \text{ where } \Gamma_m = \Gamma_m(A(t),t).$$

The cost-effectiveness condition is written:

$$\frac{\partial \Gamma_{mt}}{\partial A}\frac{\partial A}{\partial t} + \frac{\partial \Gamma_{mt}}{\partial t} = r\Gamma_m$$

Exogenous technical progress means that the marginal abatement cost decreases with time: $\partial \Gamma_{mt}/\partial t < 0$. Furthermore, the marginal abatement cost is a decreasing function of the level of abatement: $\partial \Gamma_{mt}/\partial A < 0$. It thus follows that $\partial A/\partial t < 0$. The emissions are reduced over time and the rate of decrease of emissions is increased by technical progress.

With respect to the exogenous technical progress under certainty conditions, the optimal sequential decision rule leads to reduced abatement efforts in the present, in the hope that new higher-performing technologies will enhance abatement effectiveness in the future. As a result, the initial carbon value is lower and then goes on to rise more swiftly, as exogenous technical progress accelerates the pace of abatement.

However, the introduction of uncertainty into the dynamic of climate change, the cost of innovative techniques and the risks relating to their performance substantially alters these comforting conclusions (Golier *et al.*, 2008).

Uncertainty about the scale of climate change pertains to the pace of GHG concentration, the sensitivity of rising temperatures to this concentration, the thresholds that trigger divergent feedback effects, the scope, variety and catastrophic runaway of the damage caused by temperature rise, and particularly their geographical distribution.

Uncertainty about technical progress is of two kinds: radical uncertainty concerning perfectly clean "backstop" technologies that fundamental research has yet to discover (thermonuclear energy, fuel cells) or which are still far from the market (carbon sequestration); and uncertainty about the diffusion of operational but costly techniques whose commissioning depends on adequate incentives and thus a socio-political regulation to steer and finance investment in innovation. For technical progress is endogenous. Its rhythm and direction depend on upstream investment to kick-start its mechanisms, which include learning, R&D and competition to fuel its dissemination. Governments thus need to design adapted instruments: regulations, price signals, financing packages. These will depend on the choice of the social value given to carbon.

Choice is sequential, but produces a time path that is greatly modified in comparison with the results of models developed under certainty and exogenous technical progress. Endogenous growth under uncertainty calls for a massive investment effort geared to low-carbon innovations that should be implemented as early as possible and supported by setting an initially high carbon value (Aghion *et al.*, 2009). This investment needs to cover a very broad portfolio of technologies, including ones that drive incremental technical progress in the near term and those that bring radical technical progress in the long term.

The authors construct a model where a consumer good can be produced by either a dirty or clean technology. Both types of technology may be innovative, which is to say they improve productivity thanks to the innovative component incorporated into the investment. Existing dirty technologies that are mature benefit from the learning effect that feeds incremental innovation. The latter has a snowball effect. Such innovations are thus profitable and see their profits increase through productivity gains. It follows that the dirty installed-base technologies have an initial advantage and are reinforced by the irreversibility that drives their profitability. This means that the clean technologies cannot take off unless the carbon value is sufficiently high.

The carbon value time path thus needs to be different from the one produced by the cost-effectiveness model in Box 4. It is vital not to wait until T_1 for increases in the scarcity rent to finally make these fossil fuel-consuming technologies unprofitable. The risk of cata-

strophic damage to ecological balances would be too high, especially as the two types of technology would be less easily substitutable for each other. It is essential to have a sufficiently high social value of carbon at the outset in order to channel investments into clean technology. This must be supplemented by a financial mechanism that can incentivise institutional investors, *via* a sound state guarantee, so as to provide substantial funding for the initiators of innovative environmental projects (development banks, specialist venture capital, private equity).

In practical terms, how can one determine a social value of carbon that enjoys a broad enough consensus for a country or group of countries to adopt it with a view to financing innovations likely to give a technological advantage to those involved?

What is needed is a model that computes consumption growth paths, while at the same time stopping and then reversing the degradation of environmental services. These paths depend on a sequential decision process in favour of investments to reduce CO_2 emissions, incorporating emission reduction costs and the costs of environmental degradation caused by the delay in making these investments. The process is sequential because it integrates new information that will enrich knowledge about future climate change with respect to two crucial parameters that are today deeply uncertain: the sensitivity of temperature variations to the stock of greenhouse gases and the temperature threshold that triggers non-linear effects causing irreversible damage to ecological balances.

By introducing subjective probabilities into a range of values for these parameters and coupling these assumptions to a range of possible values for global growth, paths for the social utility function of several thousand scenarios can be computed with a suitable optimal control model. The social value of carbon is determined at each point along each path. *The average social value of carbon over an entire path is equal to the average value of the marginal discounted rate of emissions compared to the level of emissions for the period.*

For each set of assumptions about carbon sensitivity, about temperature thresholds that trigger non-linear effects involved in the increase of damage and about growth, it is possible to identify a path on which the social value of carbon evolves up to a point in time where collective action based on shared knowledge will become possible (2030?, 2050?). By assigning probabilities to the different scenarios (*e.g.* equiprobability), a range of values and an average value from now until 2020 can be computed for the social value of carbon. A simulation of this type, carried out by CIRED using the RESPONSE optimal control model, came up with a range of USD9 to 80 per tonne of CO_2 with a central value of USD22 in 2020 (Hourcade *et al.*, 2011). But the values that permit reduction of the concentration rate between now and 2020 compared to the 1990 level are above USD70 if implemented immediately.

The purpose of including the social value of carbon in the calculation of investments as of today is to encourage investments that, if delayed, would be too late to prevent a systemic ecological risk.

How can the social value of carbon be integrated into economic decisions? The more classical methods involve using taxation to establish a dual price as a market price. More precisely, cap and floor pricing could be intro-

duced on the pollution permit markets in order to limit disruptive fluctuations, or a carbon tax could be introduced specifically to steer pricing. But a global, or even European carbon tax, is not currently on the agenda. There remains a third possibility: a shadow price used to value investments that carry long-run returns thanks to this price. Thus valued, this new type of public good (an amount of avoided or reduced CO_2) can serve as collateral for the issuance of financial assets that are promises on a future reduction of emissions resulting from the investments funded by these assets. Pollution permit distribution would then have to be used as a lever to ensure that the market price for these permits converges towards the notional price. This is the theoretical plinth on which we can build a sustainable financial framework.

4.6. Discounting the future under radical uncertainty

In traditional cost-benefit analysis, the discount factor is the shadow price for discounting future costs and benefits. In the usual Ramsey optimal growth model, it is the amount of consumption an economic agent is willing to give up in the present to obtain one extra unit of consumption in the future. In a perfect foresight world, the associated economic discount rate is the risk-free interest rate, equal to the rate of pure time preference plus the growth rate times the elasticity of substitution in marginal utilities. The latter parameter is equivalent to relative risk aversion under uncertainty. For this reason, the risk-free interest rate subtracts from the formula under uncertainty a term that is itself a function of the variance of the probability density function in the log of consumption.

This standard piece of theory has fed the controversy on the Stern review (2007). Orthodox economists (Nordhaus, 2007) claimed that the discount rate used by Stern was much too low (he used the rate of pure time preference estimated at 0.1% for ethical reasons) and that information drawn from financial markets should be used to extract the risk-free interest rate. What is at stake in this controversy lies in the critical importance of the discount rate for Stern's conclusions on the need to act fast and invest massively to mitigate climate change.

In his Richard T. Ely lecture (2008), *The Economics of Climate Change*, Nicholas Stern in turn revisited the controversy and explained why referring to present market rates is completely erroneous. The problem society faces is not a cost-benefit analysis along a given economic path. It is the choice between very different paths under radical uncertainty. This means that it is necessary to go back to the first principles of welfare economics.

A discount rate is the proportionate rate of fall of the value of the numeraire used in the policy evaluation. If aggregate consumption is the numeraire, as it is in standard welfare functions, the social discount rate is the rate of change of the social discount factor. If the social value of consumption at time t is: $u(c)\,e^{-\delta t}$, then the social discount factor is the marginal utility: $u'(c)e^{-\delta t}$ and the proportionate rate of fall is: $\eta(\frac{\dot{c}}{c})+\delta$. In this formula, η is the elasticity of the social marginal utility of consumption with respect to consumption. δ, is the rate of pure time preference, and has nothing to do with individual time preference. It has an ethical value and is connected to society's attitude towards its future. It should be equal to 0. The only reason for which it could be slightly positive is the possibility of an ecological collapse that terminates life on Earth as we know it.

The issue of sustainable growth in the 21st century is closely intertwined with climate change. As the latter involves huge uncertainty, framing policy is not a matter of trade-offs between present and future consumption along a given path, but rather of choosing between widely differing consumption growth

paths. The social discount rate is thus itself path-dependent, it varies over time, it is different for each uncertain sequence of outcomes and, with the highly imperfect financial markets revealed by the crisis, it differs according to the aggregate considered.

If δ must be close to 0, the value of η is itself highly ethical, since it depends upon the social distribution of income considered to be viable or desirable on a sustainable growth path. In fact, η concerns the intertemporal distribution of incomes. $\eta>1$ implies a welfare-improving redistribution of income from the richest to the poorest. In highly unequal societies like the USA, where income transfers induce strong disincentives, η is likely to be <1. In the context of uncertainty, η is interpreted as the parameter of relative risk aversion in an expected utility model of individual behaviour. However, this interpretation is quite irrelevant for the radical uncertainty faced by societies as a whole. The ethical question is one of urgency.

4.6.1. Sustainable growth and climate change

Let us revert to the modelling of the social welfare function linked to the condition of sustainable growth paths shown in Box 1. This function assumes that environmental services are not perfectly substitutable with consumption. For the economic path to be sustainable, the long-term protection of the environment should be pursued as an end in itself. The diversity of discount rates, emphasized by Stern, is plainly illustrated. There is a standard discount rate defined as above, but there is also an ecological discount rate that reflects the relative price of the environment with respect to consumption. Thus, another crucial and very uncertain parameter enters the picture, σ, which measures in a highly aggregate model to what extent consumption and environmental goods are substitutable in producing social welfare.Furthermore, as forcefully claimed by Stern, in practical terms σ is endogenous. It will change over time according to the investments in climate change mitigation and adaptation undertaken along a given path. More importantly, it will be quite different from one path to another.

In their recent sound and enlightening paper, *Ecological Intuition versus Economic Reason*, Olivier Gueant, Roger Guesnerie and Jean Michel Lasry illustrate the broad diversity of social discount rates on growth paths that differ dramatically with respect to the evolution of the relative price of the environment.

They show that the future will be quite different depending on whether $\sigma>1$ or $\sigma<1$. They call the former hypothesis *moderate environmentalism* and the latter *radical environmentalism*. They show that the ecological discount rate is lower than the standard economic discount rate, particularly when the growth rate increases and the elasticity of substitution σ decreases.

If $\sigma>1$, consumption can be substituted for the environmental services produced by natural capital, which is depleted over time. Since the contribution of environmental services to GDP declines over time, the contribution of the environment to welfare diminishes and vanishes asymptotically. Even if the ecological discount rate is always inferior to the economic discount rate, the sustainable growth paths converge in the long run to the optimal growth rate in the Ramsey growth models: $\frac{r-\delta}{\eta}$ with the interest rate (r) measuring the intertemporal substitution of consumption.

If $\sigma < 1$, the substitutability between consumption and environmental services is low. The picture changes entirely because the contribution of environmental services to GDP increases over time with an increasingly high relative price. *Environmental issues become paramount in the long run for sustainable growth paths.* The ecological discount rate is the one that matters most and it converges asymptotically in the long run towards the rate of pure time preference if natural capital is preserved. If there is a steady depletion of natural capital at the rate s, the discount rate can even turn asymptotically negative at the value: $\delta - \eta s$ with $\delta \sim 0$.

The findings of this theoretical model totally vindicate Stern's warning that comparing the different economic dynamics implied by climate change means assessing policy in a very different way to standard welfare analysis. Sustainable growth paths crucially depend on the substitutability between, on the one hand, economic, human and other types of intangible capital, and natural capital, on the other hand. The lower the substitutability between economic goods and environmental goods, the more restricted the range of sustainable growth paths. The elasticity of substitution between natural capital and other types is therefore critical.

Since this elasticity is unknown, a dilemma arises. Should governments embark on a policy of massive investment in radical innovation at the expense of present and near-term consumption, in the hope that new renewable energy investments will make environmental and economic goods more substitutable in the longer term? Or should they wait in the hope of obtaining more information on climate change developments and natural capital depletion at the risk of allowing irreversible damage to occur? We showed earlier in this paper that the considerations involved in determining the social value of carbon under uncertainty conditions lead to recommendations for innovative, massive, diversified and immediate investment. This conclusion is bolstered by an ecological systemic risk analysis that provides rigorous theoretical insights on the precautionary principle.

4.6.2. Rationalising the precautionary principle

To understand the radical uncertainty behind the illustrative theoretical model above, one must describe the complex uncertainties in the chain of interactions between the chosen targets (limiting the rise in global temperature) and the policies designed to meet these targets. Martin Weitzman's convincing paper, *On Modeling and Interpreting the Economics of Catastrophic Climate Change*, helps to clarify the stakes.

When policy is pursued to reach a target, there is presumed to be a complex chain of interrelations. The policy affects the flow of greenhouse gas emissions and one then expects that the change in flows will impact GHG stock concentrations *via* the carbon cycle. To what extent and how quickly are unknowns. The link between GHG stock and temperature rise depends on climate sensitivity. Several models run by climate experts have given widely different orders of magnitude. These global climate models must be used to make Monte Carlo simulations to estimate the probability distribution of outcomes. The links between global mean temperature changes and regional climate change are even more uncertain. They are nonetheless crucial for estimating the distribution of damage that

should then be converted into economic damage which, in turn, impact the expected present value of social welfare. Applying the condition for sustainable growth that links the value of genuine saving to social welfare changes, one can then assess the impact of a given policy. By repeating the exercise, it is possible to compare policies and their associated growth paths.

This is easier said than done. The reason is that the complex chain of interactions encapsulates non-linear feedbacks that can cause catastrophic outcomes. It means that standard cost-benefit analysis is plagued with probability density functions in the reduced-form model of aggregate expected utility that display much greater tail fatness due to structural uncertainties.

Climate sensitivity is one such interaction. For instance, the target of limiting GHG concentration to 550 ppm $\sim CO_2$ gives 24% probability of a temperature increase over 4°C and 7% probability over 5°C with a global-average climate model. Some models arrive at much higher tail probabilities. Indeed, climate sensitivity may exhibit tipping points beyond which positive feedbacks produce uncontrollable runaway. If nothing is done, business as usual could entail 5% probability of a temperature increase of 5°C to 10°C. Such an outcome would destroy a large part of the life on Earth. There would be mass species extinction and ecosystem disintegration. All icefields would thaw and the level of the sea would rise by 10 metres or more, flooding the most populous regions of the world. Rainfall patterns would be completely disrupted and the drastic changes in precipitation would cause regional desertification on a grand scale. The consequence would be mass migrations and wars that would be likely to cause a dramatic decline in world population.

If, as most reasonable people now realise, business as usual is not an option for sustainable growth, what needs to be done is to design adequate policies and set a timeframe for their implementation. In his simulations, Stern has calculated that starting from today's level of 430 ppm and setting a target of stabilising GHG concentration at $\leq$ 550 ppm will cost 1 percent of world GDP per year, with the backing of effective policies and timely decisions. Delaying action until more information becomes available on the function that links temperature increase to GHG concentration might cost three to four times as much to achieve the same target.

Weitzman's rationale now becomes clear. In the present state of knowledge, the aggregate discounted welfare function has a fat-tail probability distribution of catastrophic climate change. The loss of welfare can be bounded only by a very high number, which is nothing less than the statistical value of human civilization. As argued above, this is utterly and uniquely a question of ethics.

Weitzman models the climate-sensitivity multiplier as an unknown scale factor (s), with the probability density function of future consumption being conditional on s. He shows that, when the unknown s is inferred by climate experts on the basis of past climate outcomes, the probability distribution of future consumption posterior to Bayesian knowledge on s is a Student-t function. It converges asymptotically to a fat-tail power law. In this stochastic universe, the expected discounted factor of future costs and benefits tends to infinity and the social discount rate tends to zero.

In this way, Weitzman arrives at a "dismal theorem". The probability of disaster declines polynomially on the scale of s, while the mar-

ginal utility impact of an ecological disaster increases exponentially on the scale of s. This theorem is valid for any utility function with a positive risk aversion.

The systemic risk inherent to climate change is somewhat particular. In a systemic financial crisis not all asset classes are hit by losses. High-powered money stands out against all other assets due to its liquidity and polarises the behaviour of people in search of a refuge asset. In a systemic climate crisis, there is a possibility that overall damages hit all asset classes. All components of real wealth could suffer losses across the planet.

The dismal theorem thus validates a generalised precautionary principle for situations of potentially unlimited downside risk exposure. Because it is impossible to deduce scientific knowledge about the tail of the distribution of damages from past observations, individuals are projected into the realm of subjective uncertainty, where no market mechanism can induce any rational conduct. Appropriate policies can only be produced by collective action based on a strong ethical component that attaches importance to the welfare of future generations. As the structural uncertainty involved is obscured by high inertia, the catastrophic consequences of climate change may unfold over a timescale spanning centuries, whereas the policies needed to avert them need to be implemented today. The collective decisions are critically sensitive to the discount rate proposed. The underlying choice of the social discount rate and the flanking structural policies designed to enhance substitutability between asset classes, so as to sustain a viable growth path, amount to a generalised dynamic asset allocation strategy, legitimised by society as a whole throughout the entire world. Representative elective democracy alone cannot organise the social debate needed to legitimise the long-run vision that can shape the relevant policies. Ethics must come to bear on the choice and reproduction of elites. The criteria of social merit and recognition must change entirely. The concept of wealth needs to be radically changed in order to measure capital as a whole, and this measure must be taken into account for collective choices. A deeper, more socially responsible, participative democracy must come into being.

References

Aghion, P., D. Hemous and R. Veugelers (2009), "No Green Growth without Innovation", *Bruegel Policy Brief*, N°. 7, November.

Arrow, K. (1995), "Intergenerational Equity and the Rate of Discount in Long-term Social Investment", Stanford University Department of Economics, Working paper N°. 97-005.

Arrow, K., P. Dasgupta, L. Goulder, G. Daily, P. Ehrlich, G. Heal, S. Levin, K.-G. Mäler, S. Schneider, D. Starret and B. Walker (2004), "Are we consuming too much?", *Journal of Economic Perspectives*, 18(4): 147-172.

Arrow, K., P. Dasgupta, L. Goulder, K. Mumford and K. Oleson (2007), "China, the US and Sustainability: Perspectives Based on Comprehensive Wealth", Stanford Center for International Development, January.

Aukrust, O. (1966), "An axiomatic approach to national accounting. An outline", *Review of Income and Wealth*, 179-190, September.

Barnett, G. E. (1936), *Two Tracts by Gregory King*, John Hopkins University Press, Baltimore.

Becker, G. (1964), *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*, Columbia University Press, New York.

Beckerman, W. and H. Cameron (2007), Ethics of the discount rate in the Stern Review on the economics of climate change", *World Economics*, 8(1): 187-210.

Bénard, J. (1972), *Comptabilité nationale et modèles de politique économique*, Thémis, PUF, Paris.

Bosetti, V., C. Carraro, R. Duval, A. Sgobbi and M. Tavoni (2009), "The Role of R&D and Technological Diffusion in Climate Change Mitigation: New Perspectives Using the Witch Model", *OECD Working Paper*, N°. 664, February.

Bowman, M.J. (1968), "Principles in the Valuation of Human Capital", *Review of Income and Wealth*, 217-246, September.

Carson, C. (1975), "The History of the US National Income and Product Accounts: the Development of an Analytical Tool", *Review of Income and Wealth*, 153-181, June.

Clark, C. (1940), *The Conditions of Economic Progress*, Macmillan, London.

Corrado, A., C. Hulten and D. Sichel (2005), "Measuring Capital and Technology: an Expanded Framework", in "Measuring Capital in The New Economy", C. Corrado, J. Haltiwanger and D. Sichel (eds), *Studies in Income and Wealth*, 65: 11-41, University of Chicago Press, Chicago.

Corrado, A., C. Hulten and D. Sichel (2006), "Intangible Capital and Economic Growth", *NBER Working Paper*, N°.11948, January.

De Perthuis, C. (2010), *Et pour quelques degrés de plus*, Pearson (2nd ed.), Paris.

Desaigues, B. and P. Point (1993), *Economie du patrimoine naturel, la valorisation des bénéfices de protection de l'environnement*, Economica, Paris.

Diamond, J. (2005), *Collapse*, Penguin Books, London.

El Serafy, S. (1989), "The Proper Calculation of Income from Depletable Natural Resources", in Y. Ahmad, S. El Serafy and E. Lutz (eds), *Environmental Accounting for Sustainable Development*, World Bank, Washington, D.C.

Gilbert, M. and I. Kravis (1954), *Étude comparative des produits nationaux et du pouvoir d'achat des monnaies*, OEEC, Paris.

Gilbert, M. and Associates, *Comparative National Products* and Price Levels, OEEC, Paris.

Goldsmith, R. (1962), *The National Wealth of the US in the Postwar Period*, Princeton University Press, Princeton.

Gollier, C. (2007), "La finance durable du rapport Stern", *Revue d'Économie Politique,* 117 (4): 463-473.

Gueant, O., R. Guesnerie and J.-M. Lasry (2010), "Ecological Intuition *versus* Economic Reason", paper for the international symposium at Aix en Provence on long-term financing, 27-28 May.

Guesnerie, R. (2004), "Calcul économique et développement durable", *Revue Economique*, 55(3): 363-82.

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

Harrison, A. (1989), "Introducing Natural Capital into the SNA", in Y. Ahmad, S. El Serafy and E. Lutz (eds), *Environmental Accounting for Sustainable Development*, 19-25, World Bank, Washington, D.C.

Hicks, J. (1940), "The Valuation of Social Income", *Economica*, 7: 104-24.

Hicks, J. (1942), *The Social Framework*, Oxford University Press, London.

Hicks, J. (1961), "The Measurement of Capital in Relations with the Measurement of Other Aggregates", in F.A. Lutz and D.C. Hague (eds), *The Theory of Capital*, Macmillan, London: 18-31.

Kendrick, J.W. (1976), *The Formation and Stocks of Total Capital*, Columbia University Press, New York.

Keynes, J.M. (1940), *How to Pay for the War*, Macmillan, London.

Kravis, I., A. Heston and R. Summers (1982), *World Product and Income: International Comparisons of Real Gross Products*, John Hopkins University Press, Baltimore.

Kuznets, S. (1973), *Population, Capital and Growth: Selected Essays*, Norton, New York.

Landes, D. (1998), *The Wealth and Poverty of Nations*, Little, Brown and Company, London.

Leontiev, W. (1941), *The Structure of the American Economy 1919-1929: an Empirical Application of Equilibrium Analysis*, Harvard University Press, Cambridge, MA.

Lev, B. (2001), *Intangibles: Management, Measurement and Reporting*, Brookings Institution Press, Washington, D.C.

Lutz, E. (ed.) (1993), *Towards Imported Accounting for the Environment*, World Bank, Washington, D.C.

Maddison, A. (2003), *The World Economy: Historical Statistics*, Development Center Studies, OECD, Paris.

Maddison, A. (2007), *Contours of the World Economy 1-2030 AD*, in Essays in Macro-Economic History, Chap. 6, Oxford University Press, Oxford.

Mairesse, J. (1972), *L'évaluation du capital fixe productif - Méthodes et résultats*, Collections de l'INSEE, Series C, N°. 18-19, November.

Mairesse, J. and Y. Kosoglu (2005), "Issues in Measuring Knowledge: the Contribution of R&D and ICT to Growth", communication to the Advancing Knowledge and Knowledge Economies, Conference at the National Academies, Washington, D.C., 10-11 January.

Meade, J. and R. Stone (1941), "The Construction of Tables of National Income, Expenditure, Savings and Investment ", *The Economic Journal*, 51 (1941): 216-31.

Nordhaus, W. and J. Tobin (1973), "Is Growth Obsolete?", in M. Moss (ed.); *The Measurement of Economic and Social Performance*, Columbia University Press, New York.

Nordhaus, W. (2007), "A Review of the Stern Review on the Economics of Climate Change", *Journal of Economic Literature*, 45(3): 686-702.

Parson, E. (2007), "The Big One: a Review of Richard Posner's Catastrophe: Risk and Response, *Journal of Economic Literature*, 45: 147-164.

Pearce, D .W. and G. Atkinson (1993), "Capital theory and the measurement of sustainable development: an indicator of 'weak' sustainability", *Ecological Economics*, 8: 103-108.

Pigou, A.C. (1932), *The Economics of Welfare*, (4th ed.), Macmillan, London.

Posner R. (2004), *Catastrophe: Risk and Response*, Oxford University Press, Oxford.

Ramsey, F.P. (1928), "A Mathematical Theory of Saving", *Economic Journal*, 38(152): 543-559.

Sametz, A.W. (1968), "Production of Goods and Services. The Measurement of Economic Growth", in E.B. Sheldon and W. Moore (eds), *Indicators of Social Change. Concepts and Measurement*, 77-96, Russel Sage Foundation, New York.

Samuelson, P. (1950), "Evaluation of real national income", *Oxford Economic Papers*, 2: 1-29.

Schultz, T.W. (1961), "Investment in Human Capital", *American Economic Review*, 51(1): 1-17.

Sen, A. (1976), "Real National Income", *Review of Economic Studies*, 43(1): 19-39.

Sen, A. (1979), "The welfare basis of income comparisons: a survey", *Journal of Economic Literature*, 17(1): 1–45.

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

Stern, N. (2007), *The Economics of Climate Change. The Stern Review*, Cambridge University Press, Cambridge.

Stern, N. (2008), "The Economics of Climate Change", Richard T. Ely Lecture, *American Economic Review Papers and Proceedings*, 98(2): 1-37.

Stone, R. (1947), "Definition and Measurement of the National Income and Related totals", in *Measurement of national Income and the Construction of Social Accounts*, UN.

Stone, R. (1956), *Quantity and price Indexes in National Accounts*, OEEC, Paris.

Studensky, P. (1958), *The Income of Nations. Theory, Measurement, and Analysis: Past and Present*, New York University Press, New York.

Uno, K. and P. Bartelemus (eds) (1998), *Environmental Accounting in Theory and Practice*, Kluwer Academic, Boston.

Vanoli, A. (1995), "Reflections on Environmental Accounting Issues", *Review of Income and Wealth*, 41: 113–137.

Vanoli, A. (2002), *Une Histoire de la Comptabilité Nationale*, Repères, La Découverte, Paris.

Weitzman, M. (2009), "On Modeling and Interpreting the Economics of Catastrophic Climate Change", *Review of Economics and Statistics*, 91(1): 1-19 .

World Bank (1995), *Monitoring Environmental Progress*, The World Bank, Washington, D.C.

World Bank (2006), *Where is the Wealth of Nations? Measuring Capital for the 21st Century*, The World Bank, Washington, D.C.

Comment: "Sustainable Growth: Do We Really Measure the Challenge?"

Frank Lecocq, AgroParisTech

Introduction

Michel Aglietta's contribution offers a very rich and insightful picture of the history, state-of-the-art, and challenges associated with incorporating sustainability considerations in macro measurement. Building on a brief history of national account systems (Introduction and section 4.1.), the paper calls for two key improvements to embark sustainability considerations: (i) broadening the concept of capital to total wealth, which encompasses, *inter alia*, human, social and natural capital; and (ii) adopting a sustainability condition, which states that the first derivative of society's intertemporal welfare function must be non-negative at any point in time. Linking (i) and (ii) yields the condition that a development path is sustainable if, at every date, *genuine savings* are non-negative – a condition valid under broad assumptions (section 4.2.).

Acknowledging that sustainable policies are likely to require significant public investment, section 4.3. examines the long-term sustainability of public finances. It calls for measuring solvency based on a broader indicator than debt/GDP ratio, namely the net worth of the public sector, which takes into account, *inter alia*, future contingent liabilities linked to pensions. The paper then turns to the extended concept of capital, examining successively the issues associated with the measurement of intangibles (section 4.4.) and of natural capital, including the social value of carbon (section 4.5.). Section 4.6. concludes by discussing two parameters critical for measuring the long-term implications of environmental policies: the elasticity of substitution between environmental services and consumption on the one hand, and the discount rate on the other.

As should be obvious from the above summary, Michel Aglietta's paper covers a lot of ground and, consequently, opens up many interesting discussions. The present note focuses solely on two points of particular relevance for the conference: the use of national account systems for the evaluation of policies with implications in the very long term, such as climate mitigation policies, and the impact of improving macro measurement techniques on effective policy decisions.

How can improved national accounts inform long-term policies? A first major theme in the

paper is that improved national account systems could and should be used to guide policies that have implications in the very long term, such as climate mitigation policies. By very long term, we mean here at least 5-10 decades in the future (though mitigation policies will also have implications earlier), a far longer time horizon than in most economic sectors, barring transportation infrastructure, housing or forestry.

Over such a time horizon, the sustainability condition that net savings be positive *at each point in time* seems restrictive. Precisely, if the improved accounting system is comprehensive, *i.e.*, if total wealth is correctly captured, then there can be no gain associated with negative genuine savings. The condition is valid. But if in fact some aspects of total wealth are not captured in the improved system, then imposing the sustainability condition at each point in time might be too restrictive. For example, investments in ecosystem "quality", *e.g.*, to improve the functioning or the resilience of existing ecosystems, might be difficult to capture (at least within the definition of natural capital discussed in section 4.5.), yet be beneficial to society (much like an environmental intangibles).

Over such a long time period, also, hardly any economic parameter can be considered constant. And, as the paper points out, most economic parameters are in fact endogenous to the set of policies selected in the first place. In other words, evaluating different sets of policies related to climate change (*e.g.*, a policy that stabilizes greenhouse gases concentration at a given level in the atmosphere, *versus* a reference scenario in which no major mitigation policy is implemented) amounts to comparing economies that rapidly become very different in terms of stock and composition of capital (artificial and natural), in terms of technology, in terms of size, location and distribution of income of the population, and possibly in terms of preferences.

In this context, the assumptions under which the sustainability criteria have been derived may no longer be valid. Precisely, the sustainability criteria rest on the assumption that the rules governing the allocation of output remain constant over time (section 4.2., Box 1). Yet this condition may not hold when policies leading to very different outcomes in the long run are compared. In such case, the intertemporal welfare V becomes an explicit function of time, no longer solely a function of capital stocks at *t*. How this impacts the sustainability condition should be explored.

Even assuming the above problem is solved, differences in economic outcomes might make the ranking between alternatives difficult to establish, even within an improved accounting system. For example, if people are conscious of, and derive disutility from, the animal and vegetal species that disappear during their lifetime, but are indifferent to species that disappeared before they were born, then the same level of utility might be attained, in the long run, with any level of biodiversity. Does this imply that we should not engage in long-term biodiversity protection programs? Certainly not. But the argument for that cannot be derived from a pure welfare measurement.

Improved national accounts as preconditions for more sustainable policies

A second key message of the paper is that extending national account systems is a necessary condition for the adoption of more

sustainable policies. It seems very reasonable to assume that using genuine savings can be a powerful argument to shed light on the inefficiencies of current public policies, and to point to more sustainable solutions. On the other hand, it is not obvious why changing the measurement tool would be a necessary "precondition for designing adequate investment policies". Nor is it clear why it would be sufficient, as even "quasi Pareto-optimal policies" are often not adopted (Stiglitz, 1998). The key point here is that it would be very useful to understand the political economy of *how* an improved accounting system might be most effective in influencing national and global policies, since this might in turn have a strong bearing on the design of such improved tool and on the selection of priorities for improvement.

The following example tries to illustrate how an improved accounting system might help in designing climate policies. In most climate change analysis, mitigation (or adaptation, or combinations thereof) is measured relative to a reference scenario in which no climate policy is implemented. In fact, it is common to use a reference in which climate change does not exist at all. Yet this is a dream world: without action against climate change, societies will incur the full costs of climate damage (net of reactive adaptation) (Shalizi and Lecocq, 2010).

A key reason why a "without-climate-change" reference is often used is that the impacts of climate change on growth remain difficult to quantify. An improved version of national accounts would make it much easier to construct a correct reference, if only because the impacts of climate change on natural capital (*e.g.*, on forest resources, water cycles, or ocean biodiversity, etc.) could be captured directly. Of course, this should not deter us from trying to assess climate change implications for GDP, because implications for markets are important per se; and because understanding the dynamics by which climate change impacts economic activity is also critical to finding relevant adaptation policies. Yet it would constitute a notable improvement over current analysis of climate policies

Conclusion

National accounting is generally not the first concept that springs to mind when discussing sustainability issues. Yet Michel Aglietta's paper makes a convincing case that the two are linked; namely that current national accounts systems are lacking, that a major overhaul is a necessary condition for more sustainable policies to be adopted, and that solid theoretical foundations exist to do so. As such, this is a very insightful and thought-provoking paper, of which I strongly recommend the reading.

References

Shalizi, Z. and F. Lecocq, (2010), "To Mitigate or to Adapt: Is that the Question? Observations on an Appropriate Response to the Climate Change Challenge to Development Strategies", *World Bank Research Observer* 25(2): 295–321.

Stiglitz, J. (1998), "Distinguished Lecture on Economics in Government. The Private Uses of Public Interests: Incentives and Institutions", *Journal of Economic Perspectives* 12(2): 3–22.

Conférences & Séminaires

Acronyms and Abbreviations

Conférences & Séminaires

List of Acronyms and Abbreviations

AFD	Agence Française de Développement
BHPS	British Household Panel Survey
CEPREMAP	Centre pour la recherche économique et ses applications
CHIP	Chinese Household Income Project
CIRED	Centre international de recherche sur l'environnement et le développement
CPI	Consumer Price Index
DHS	Demographic and Health Survey
EHCP	European Community Household Panel
ESS	European Social Survey
EUDN	European Development Research Network
GDP	Gross Domestic Product
GHG	Greenhouse Gas
GSOEP	German Socio-Economic Panel
GSS	General Social Survey
FAO	Food and Agriculture Organization
HDI	Human Development Index
HILDA	Household Income and Labor Dynamics in Australia
IARIW	International Association for Research in Income and Wealth
ICP	International Comparison Project
IFPRI	International Food Policy Research Institute
INRA	Institut national de la recherche agronomique
IPCC	Intergovernmental Panel on Climate Change
ISSP	International Social Survey Programme
LITS	Life in Transition Survey
LSMS	Living Standard Measurement Survey
MDGs	Millenium Development Goals
MICS	Multiple Indicator Cluster Survey
NI	National Income
NP	National Product
MPI	Multidimensional Poverty Index
NFHS	National Family Health Survey
OECD	Organisation for Economic Co-operation and Development
OEEC	Organisation for European Economic Cooperation

OPHI	Oxford Poverty and Human Development Initiative
PPP	Purchasing Power Parity
PWT	Penn World Table
REDD	Reduced Emissions from Deforestation and Forest Degradation
RLMS	Russian Longitudinal Monitoring Survey
SALDRU	South Africa Labor Development Research Unit
SNA	System of National Accounts
SWB	Self-reported Well-being
TFP	Total Factor Productivity
UNDP	United Nations Development Programme
ULMS	Ukrainian Longitudinal Monitoring Survey
WDI	World Development Indicators
WDR	World Development Report
WHO	World Health Organization
WHS	World Health Survey
WVS	World Values Survey

Conférences
& Séminaires

Authors' Biographies

Authors' Biographies

Michel AGLIETTA. Emeritus Professor at the University of Paris-X Nanterre, scientific advisor for the Center of International Studies and Forecasting (CEPII), and for Groupama AM. He is a former member of the French Council of Economic Advisors attached to the Prime Minister, and former senior member of the *Institut Universitaire de France*. His most recent books are *Corporate Governance adrift*, with A. Rebérioux, Edward Elgar (ed.), 2005; *Désordres dans le capitalisme mondial* with A. Berrebi, Odile Jacob, 2007; *Macroéconomie Financière*, 5th ed. entirely revised, collection Grands Repères, La Découverte, 2008; La Crise, Michalon, November 2008; *Crise et rénovation de la Finance,* with Sandra Rigot, Odile Jacob, March 2009; *Les hedge funds: entrepreneurs ou requins de la finance?,* with Sabrina Khanniche and Sandra Rigot, Perrin, January 2010 and *La crise : les voies de sortie*, ed. Michalon, May 2010.

Sabina ALKIRE directs the Oxford Poverty & Human Development Initiative (OPHI), a research centre in the Oxford Department of International Development, Queen Elizabeth House, University of Oxford. Her publications focus on the capability approach, multidimensional poverty measurement, agency, and human development. She has worked previously at Harvard University, the Sen-Ogata Commission on Human Security at the United Nations, the World Bank, and Oxfam. She holds a D.Phil. in Economics from Magdalen College, Oxford.

François BOURGUIGNON is the Director of the Paris School of Economics. After four years as the Chief Economist and first Vice President of the World Bank in Washington, he returned to France in 2007, where he took up his former position of Professor of Economics at the EHESS (school for advanced studies in the social sciences). Trained as a statistician, he obtained a Ph.D. in Economics at the University of Western Ontario, followed by a State Doctorate at the University of Orleans in France. His work is both theoretical and empirical and principally focuses on the distribution and redistribution of revenue in developing and developed countries. He is the author of a great number of books and papers in specialized national and international economic journals. He has taught in universities all over the world and, during his career, has received a number of scientific distinctions. On account of his broad experience, he is often asked to advise Governments and international organisations.

Andrew CLARK holds a Ph.D. from the London School of Economics. He is currently a CNRS Research Professor at the Paris School of Economics (PSE), and previously held posts at Dartmouth, Essex, CEPREMAP, DELTA, the OECD and the University of Orleans. His work has largely focussed on the interface between psychology, sociology and economics; in particular, using job and life satisfaction scores, and other psychological indices, such as proxy measures of utility. The broad area is social interactions and social learning. One research field has

been that of relative utility or comparisons (to others like you, to others in the same household, and to yourself in the past), finding evidence of such comparisons with respect to both income and unemployment. This work has spilled over into theoretical and empirical work on evidence for and the implications of following behaviour and learning from others' actions. Recent work has involved collaboration with psychologists to map out habituation to life events (such as job loss, marriage, and divorce) using long-run panel data. In addition, direct measures of utility allow direct tests of popular models of the labour market. In this spirit, his work has looked at unemployment, quits, and labour market rents.

Angus DEATON is the Dwight D. Eisenhower Professor of Economics and International Affairs at Princeton's Woodrow Wilson School and Department of Economics. He holds B.A., M.A. and Ph.D. degrees from the University of Cambridge. His interests include health, development, poverty, and inequality; he is the author of four books and many papers. In 2006, he chaired a panel charged with the evaluation of World Bank research over the previous decade. He has served on National Academy panels on poverty and family assistance and on price and cost-of-living index numbers. He is a Senior Research Scientist with the Gallup Organization, working on their World Poll, and exploring the global links between life evaluation, hedonic well-being, income and health. He was the first recipient of the Econometric Society's Frisch Medal, and edited *Econometrica* from 1984–1988. He is a Fellow of the British Academy, of the Royal Society of Edinburgh, of the Econometric Society and of the American Academy of Arts and Sciences. He was President of the American Economic Association in 2009. He holds honorary degrees from the University of Rome, Tor Vergata, University College London, and the University of St. Andrews, and is an honorary fellow of Fitzwilliam College, Cambridge.

Claudia SENIK is Professor of Economics at the Paris School of Economics and at the University Paris-Sorbonne. She is also a member of the IZA and of the *Institut Universitaire de France*. Educated at the *Ecole Normale Supérieure*, she received her Ph.D. from the EHESS. Her main research areas include public economics and happiness studies, with a special interest in the subjective welfare effects of income growth and income distribution. She is the author of a large number of papers, published in international reviews such as the *Review of Economics and Statistics, the Journal of Public Economics, the Economic Journal, the Journal of Economic Behavior and Organization, the Journal of Comparative Economics, the Economics of Transition, Economica, Economic Policy*, etc. She is currently responsible for a CEPREMAP research program dedicated to *Income and Happiness* (with Andrew Clark), for the *Economics and Psychology* seminar (with Guillaume Hollard) at the Paris School of Economics, and for the CEPREMAP Public Policy and Redistribution program (with Karine van der Straeten and Brigitte Dormont).

Discussants

Pramila KRISHNAN is a Senior Lecturer at the Faculty of Economics at the University of Cambridge and a Fellow of Jesus College, Cambridge. She began her academic life as an econometrician working on models of self-selection before seeing the light and moving on to work in development economics. Her research has concentrated on applied microeconomics and she has worked on topics ranging from household portfolios of poor households, to risk-sharing and informal insurance, social networks, non-cognitive skills and whether migrants might be both rich and happy.

Franck LECOCQ is an economist with the AgroParisTech (the Paris Institute of Technology for Life, Food and Environmental Sciences). He is senior researcher at CIRED, France, and was deputy director of the Laboratory of Forestry Economics, a joint research unit of AgroParisTech and INRA (French National Institute for Agronomical Research) at the time of the conference. Franck has published extensively on the economics of climate change, and his current research focuses on the relationships between climate change and sustainable development, the economics of adaptation to climate change, and the economics of the forest/carbon interface. Franck Lecocq was formerly a research fellow at CIRED, France, and at the Development Economics Research Group of the World Bank. He holds a Ph.D. in economics from AgroParisTech, and is qualified to supervise Ph.D.s from Paris-Dauphine University. He is a member of the Economic Council for Sustainable Development of the French Ministry of Environment, Energy, Sustainable Development, and a lead author of the IPCC Fourth and Fifth Assessment Reports.

Alemayehu SEYOUM TAFFESSE is currently a Research Fellow at the International Food Policy Research Institute (IFPRI). His research spans individual aspirations and well-being, impact evaluation of government programs, household risk and vulnerability, government policy and agriculture, and economic governance in Ethiopia and further afield. He has previously worked as an Assistant Professor of economics at Addis Ababa University, an Economic Affairs Officer at the United Nations Economic Commission for Africa, and the Director of the African Centre for Economic and Historical Studies (an independent research outfit in Addis Ababa that he established in 2004). Alemayehu holds a D.Phil. degree in Economics from the University of Oxford.

Conférences
& Séminaires

What is AFD ?

Agence Française de Développement is a public development finance institution that has worked to fight poverty and support economic growth in developing countries and the French Overseas Communities for 70 years. AFD executes the French government's development aid policies.

Through offices in more than fifty countries and nine French Overseas Communities, AFD provides financing and support for projects that improve people's living conditions, promote economic growth and protect the planet: schooling, maternal healthcare, help for farmers and small business owners, clean water supply, tropical forest preservation, and fighting climate change, among other concerns.

In 2010, AFD approved more than €6.8 billion for financing aid activities in developing countries and the French Overseas Communities. The funds will help 13 million children go to school, improve drinking water access for 33 million people and provide €428 million in microloans benefiting more than 700,000 people. Energy efficiency projects financed by AFD in 2010 will save nearly 5 million tons of carbon dioxide emissions annually.

www.afd.fr

Agence Française de Développement
5, rue Roland Barthes – 75598 Paris cedex 12
Tel.: 33 (1) 53 44 31 31 – **www.afd.fr**

ISSN: 2118-3872

Measure for Measure

How Well Do We Measure Development?

The Agence Française de Développement and the EUDN network of European economists have been co-organising an annual conference on development for eight years now. Over time, this event has become a major landmark in Europe for the development community. The one that took place on 1st December 2010 gathered over 1,000 participants from thirty countries. It returned to a core question: the measure of development and, thus, its very nature. Do we really know what we mean when we talk of "development" if we cannot agree on how to measure it?

Angus Deaton (University of Princeton), Sabina Alkire (University of Oxford), Claudia Senik, Andrew Clark (Paris School of Economics), and Michel Aglietta (University of Paris X) each gave their own answers by offering fresh research contributions. They thereby pointed to a set of facts and questions that give a holistic and up-to-date vision of the issue of measurement and of its implications. Francois Bourguignon (PSE), Alemayehu Seyoum Taffesse (IFPRI Addis Ababa), Pramila Krishnan (University of Cambridge) and Frank Lecocq (AgroParisTech) further enriched these analyses with their own critical thinking. The quality of all these contributions and of the ensuing debates with the public was widely applauded. This is well reflected in the present publication.

CONTACT
Robert PECCOUD
Research Department, AFD
peccoudr@afd.fr

Conférences & Séminaires