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).

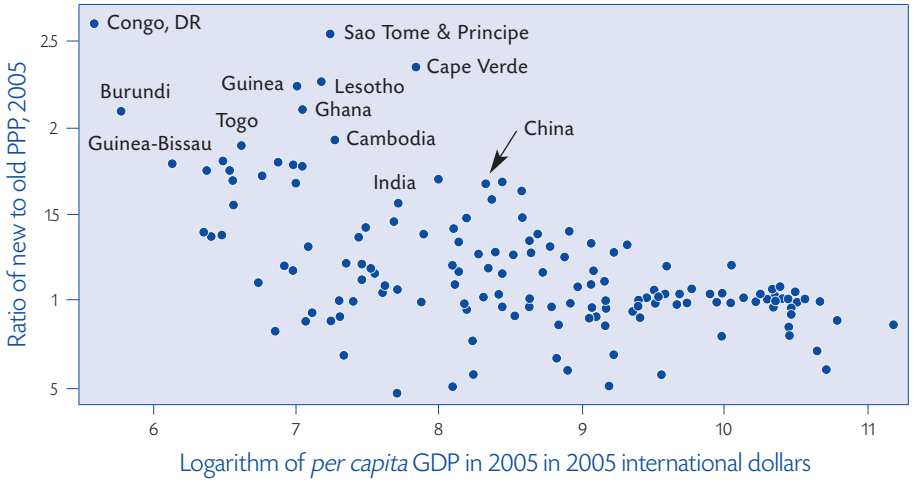
What do we conclude from all this? First, it is not obvious how to maintain a constant poverty line through a new round of the ICP. The Bank's procedures do not do so, causing large shifts in both the structure and total of world poverty. Perhaps the level of global poverty is less important than its 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

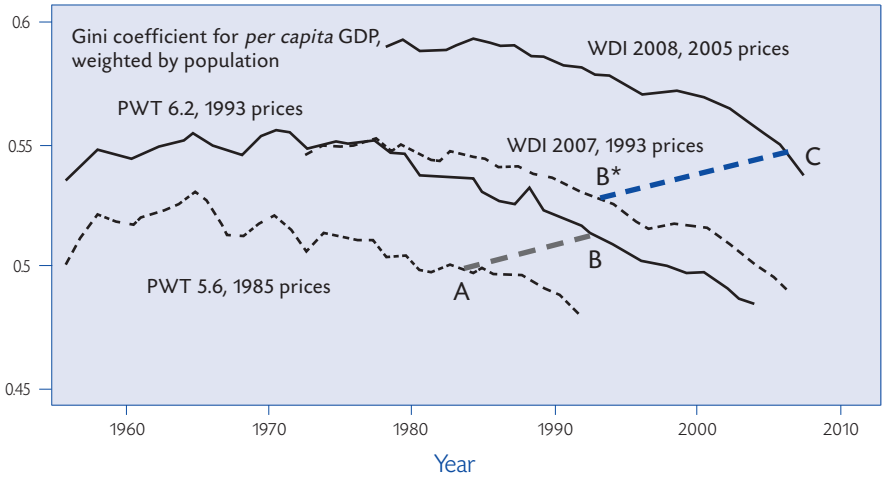


Source: author's calculations, taken from Deaton (2010).

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.

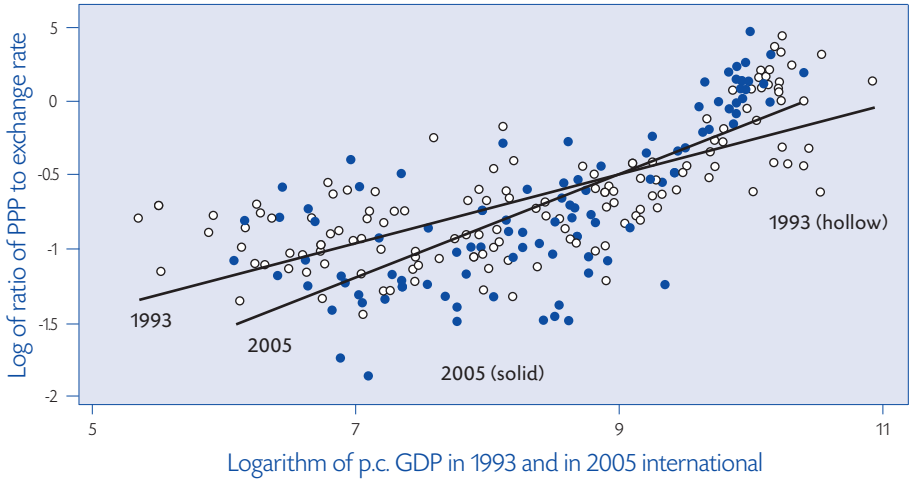
Why do ICP revisions increase the spread of national incomes? One answer is that they do not, at least in general. Each ICP revision has contained a large number of methodological improvements over previous rounds, and these revisions will certainly change measured inequality. In Figure 2, we are looking 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

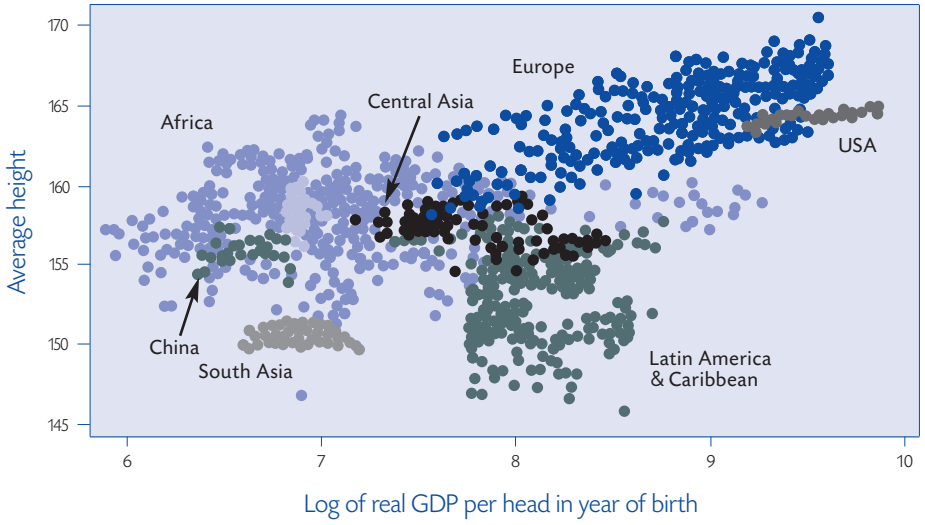
meeting the demands of disease and work, not *gross* nutrition – the intake of food – that affects physical and mental health and the growth and development of children. Of course, it is not a good thing to be hungry, or to get fewer nutrients than are needed, though the measurement of need by fixed 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).

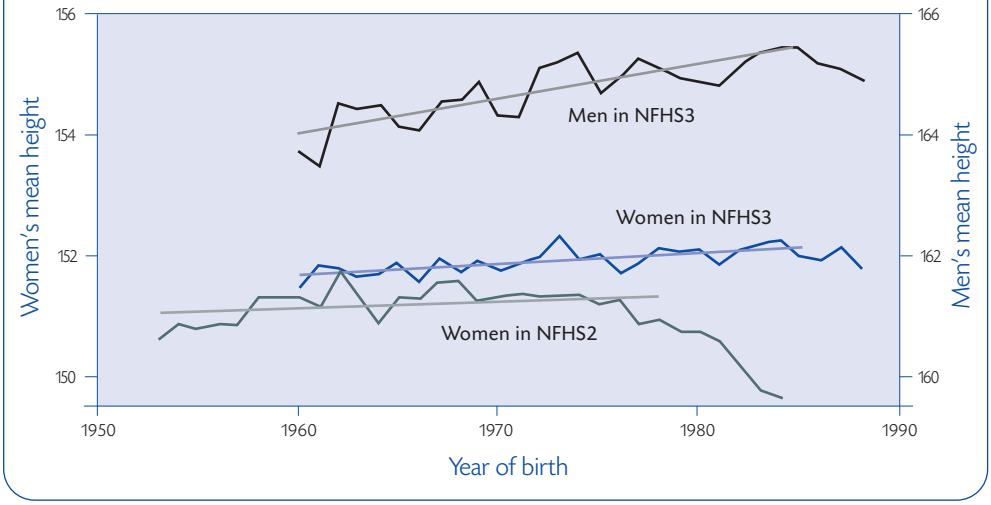
1.3.3. Calories and nutrition in India

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.