NBER WORKING PAPER SERIES

UNDERSTANDING PPPS AND PPP-BASED NATIONAL ACCOUNTS

Angus Deaton
Alan Heston

Working Paper 14499
http://www.nber.org/papers/w14499

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2008

Both authors were members of the Technical Advisory Group for the 2005 International Comparison Program, which was chaired by Heston. This paper represents their personal views, and should not be taken to represent the views of the ICP, the World Bank, or any other international institution, none of whom financially supported this work. Support from NSF grant SES-0648769-002 is gratefully acknowledged by Heston. Parts of this paper were presented at CRIW-NBER Summer Workshop, July 14-15, 2008. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Understanding PPPs and PPP-based national accounts
Angus Deaton and Alan Heston
NBER Working Paper No. 14499
November 2008
JEL No. E01,N1,O47

ABSTRACT

PPP-based national accounts have become an important part of the database for macroeconomists, development economists, and economic historians. Frequently used global data come from the Penn World Table (PWT) and the World Bank's World Development Indicators; a substantial fraction of the world is also covered in the PPP accounts produced by the OECD and the European Union. This paper provides an overview of how these data are constructed, and discusses both the theory and the practical problems of implementing it. All of these data are underpinned by the International Comparison Program (ICP), which collects data on prices worldwide. The most recent round of the ICP was for 2005 with final results published in early 2008; version 7.0 of the Penn World Table will soon incorporate these results. The 2005 ICP, like earlier rounds, involved substantial revisions to previous data, most notably revising downwards the size of the Chinese (40 percent smaller) and Indian (36 percent) economies. We discuss the reasons for the revisions, and assess their plausibility. We focus on four important areas: how to handle international differences in quality, the treatment of urban and rural areas of large countries such as China, India, and Brazil, how to estimate prices for government services, health, and education, and the effects of the regional structure of the ICP. All of these affect the interpretation of previous data, as well as the current revisions. We discuss previous revisions of the PWT, and their effects on various kinds of econometric analysis. The paper concludes with health warnings that should be kept in mind when using these data, which are not always suitable for the purposes to which they are put. Some international comparisons are close to impossible, even in theory, and in others, the practical difficulties make comparison exceedingly hazardous.

Angus Deaton
328 Wallace Hall
Woodrow Wilson School
Princeton University
Princeton, NJ 08544-1013
and NBER
deaton@princeton.edu

Alan Heston
433 12th St. SE
Washington DC 20003
aheston@econ.upenn.edu

1. Introduction

Over the last thirty years, more and better data have fundamentally changed the practice of both microeconomics and macroeconomics. No new data have been more important and more influential than those from the International Comparison Program (ICP). By the late 1960s, the theory of economic growth that had begun with Solow's great paper had become a largely theoretical enterprise. But by the late 1980s, the Penn World Table had evolved from a small set of illustrative calculations to a multi-country panel big enough for econometric analysis, particularly Mark 5 which contained up to 39 years of data on 138 countries, Summers and Heston (1991). These data helped bring about a new growth economics, with theoretical developments consistently related to evidence. There has been a huge explosion of work since then, trying to understand the mechanics of growth, linking growth and politics, and forging an integration of macroeconomics, economic development, and economic history, the last supported by the companion creation of long-run historical data by Maddison (2003). Purchasing power parity data provide a common measuring rod that allows comparison, not only India and America now, but India now with Britain before the industrial revolution. It allows the World Bank to estimate the number of poor in the world, and permits plausible conjectures about which places and which times have seen the greatest riches and the greatest poverty in human history, Pritchett (1997).

Although the international data are widely used, the way that they are constructed is not always as widely understood. Nor is it easy to find out why different, commonly used sources give different estimates. Perhaps the most familiar of these sources are the PPP data given in the PWT, on the one hand, and in the World Bank's World Development

Indicators (WDI) on the other. Eurostat and the OECD currently cover 55 countries in the OECD, Europe, and the Confederation of Independent States (CIS), and publish time-series data back to 1980 for a subset of those countries. This paper aims to provide an overview of the most recent round of data collection, together with the underlying economic theory, an explanation of why different sources give different numbers, and some health warnings for their use. We pay most attention to the PWT, if only because it is the only source that gives long time series for a disaggregated set of national accounts There is an enormous amount of detail that goes into the collection of the data and the construction of the accounts, but our aim here is to focus on a few key issues that are likely to be important to practicing macroeconomists, and where we think more knowledge is likely to be useful in practice.

An overview of the most recent round of international price comparisons is provided in the final report of the ICP, World Bank (2008a), and the details can be found online in the ICP Handbook, World Bank (2008b). Our account here starts with the price-index theory that underlies the international accounts, and emphasizes the differences between multilateral and standard price comparisons. This also allows us to explain why the same set of underlying prices can give rise to several sets of apparently inconsistent national accounts. We then discuss some of the practical difficulties. Among the most important of these are how to handle international differences in the quality of goods and services across countries, how to price "comparison resistant" items such as government services or the imputed rents of owner occupiers, and how to calculate accounts for countries and periods that are not covered by the underlying data collection. There are also specific "health warnings" with any specific set of accounts; in the latest round the most

important concern the regional structure of the data collection, the balance between urban and rural data collection, and the role of India, China, and the countries of the former Soviet Union.

Most economists who use PPP data would currently work with the latest version of the Penn World Table, Version 6.2, which does not include the data from the latest ICP round for 2005. Version 7.0 of the PWT, which is currently in preparation, will incorporate these data, and will adjust the past data to provide a coherent set of numbers as far back as 1950. So we emphasize some of the areas where the 2005 round is different from earlier rounds, in part because these will cause substantial revisions compared with Version 6.2, but also to illustrate the changes that have taken place in the past with every new round of data collection. These revisions are often substantial. For example, the 2007 version of the World Development Indicators (WDI), World Bank (2007), lists 2005 per capita GDP for China as $6,757 and for India as $3,452, both in current international dollars. The 2008 version, World Bank (2008c), which includes the new ICP data, gives, for the same year, and the same concept $4,088 for China and $2,222 for India. For comparison, GDP per capita at market exchange rates is $1,721 for China and $797 for India. We shall have something to say about what drives these revisions but it is hard not to speculate about which previously established econometric results survive the incorporation of these revisions into the PWT.

Economists are most familiar with PPP accounts through the PWT and, secondarily, through Maddison's data. The underlying data all come from one or more rounds of the ICP, which started as a joint UNSO-Penn project, then a UNSO project, and most recently a World Bank managed project endorsed by the UN Statistical Commission .

The main business of the ICP is to collect data on the prices of thousands of comparable goods and services in many countries, 148 in the 2005 round. These prices, together with the national accounts for each country, are used to construct a set of price indexes that compare, for example, the price of consumption or investment in India relative to the price of consumption or investment in the US, expressed in rupees per dollar. "Volume" estimates—in current prices—come from dividing expenditures by the price indexes, in the example, giving estimates of both Indian and US consumption or investment in current international dollars. Several preliminary points follow from this structure.

First, there are many different reasonable formulas for price indexes, and these will give rise to different indexes and different sets of accounts, which is (one of the reasons) why the GDP estimates published by the World Bank in the WDI are different from those that appear in the PWT. Which of these is most appropriate depends on the purposes to which the data are to be put; as is the case with most index number questions, there is typically no unique right answer. These issues are familiar from standard within-country accounting where, for example, Paasche and Laspeyres indexes are not identical. But the differences tend to be more important in international comparisons over space than in national comparisons over time, because price index weights are much more different between India and the US, for example, than within the US or within India a few years apart.

Second, neither the ICP itself, nor the PWT or the WDI, collect or edit the expenditure data, which come directly from the national accounts of the countries. (Recent rounds of the ICP have aimed to help improve statistical capacity in participating countries, but did not construct domestic national accounts.) The national income

accounts of many low-income countries are very weak, with procedures that have sometimes not been updated for decades, see in particular Seers (1983, 18–27) who describes many of the problem areas. Seers notes that no one has ever taken up his standing $10,000 wager that, working entirely within the rules set by the SNA, he could obtain any growth rate for any country, and that his estimates would pass professional scrutiny. In many cases, the prices collected under the auspices of the ICP may be more accurate than the GDP numbers with which they are combined.

Third, it is important to understand that aspects of the exercise are close to being impossible in theory, and are therefore not amenable to data improvement. Making price comparisons between Canada and the US, or between the countries of the OECD, is relatively straightforward because the same goods and services are widely available in all. But when it comes to comparing a rural Thai agricultural laborer, who lives almost entirely on rice, with his Ethiopian counterpart, who lives on *teff*, we have no basis for comparison. Rice is hard to find in Ethiopia and *teff* is impossible to find in Thailand, so price comparisons are not possible. This is an extreme case, but many goods and services that are widely consumed in rich countries are not available at all in poor countries, or are only available at high-priced stories in a few large cities. One general rule is that the comparisons become less reliable the further apart are the structures of GDP (or its components) of the countries being compared. This is essentially the same phenomenon as the increasing unreliability of long-run historical comparisons the further back we go.

With these caveats in mind, we plunge into the formulas, which assume that prices and expenditures are available for all countries on a common set of goods and services. We shall return to the reality in Section 3.

2. Multilateral price indexes in theory

In a world in which the law of one price were true, market exchange rates would be all that we would need for converting accounts in one currency into another. The price of any item in one country would be the price in any other converted at the exchange rate, and the same would be true for a price index for consumption, investment, GDP or whatever. For various reasons, see Rogoff (1996) for a review, relative prices are different in different countries, so that it is useful to compare prices directly, and to calculate price indexes for GDP and its components. Given a set of prices of all the goods and services in consumption (say, and we use consumption as an example), it is straightforward to use standard formulas (Paasche, Laspeyres, or Fisher, for example) to compute consumer price indexes for any one country in terms of any other. But these "binary" indexes do not give us what we want. If there are M countries, they yield an M by M matrix of price indexes, not a vector of M price indexes, one for each country which, like market exchange rates, would allow us a unique way of converting the price level of one country into another. More formally, we require that the matrix of the price indexes have two properties, first that the price index of country c in terms of country d should be the reciprocal of the price index of d in terms of c, and second, that the indexes be transitive, so that the price of Thailand with Botswana as base is the same whether computed in one step, or computed in two or many steps, from Thailand to Peru and then from Peru to Botswana, or through other intermediate countries. It is straightforward to show that these requirements are satisfied if, and only if, there exists a vector of M price

indexes such that element *c*, *d* of the matrix is given by the ratio of the price index for *d* to that of *c*.

Perhaps the obvious approach to economists—though not to national income statisticians—is to work with cost-of-living indexes which, in theory, provide the price indexes that we need. If prices in country *c* are written as an *n*-vector p^c, and if tastes are homothetic and identical in all countries, the cost or expenditure function can be written as $u^c\alpha(p^c)$ for utility level u^c and some linearly homogeneous function $\alpha(.)$ which is not indexed on *c*. The α functions immediately give us the price indexes that we need, so that if we (arbitrarily) take country 1 as numeraire (the US in all of these calculations), the consumption PPP for *c* in international dollars is simply

$$P^c = \alpha(p^c)/\alpha(p^1). \qquad (1)$$

These PPPs make no assumption about relative prices being the same in all countries; essentially the assumption of identical homothetic tastes replaces identical relative prices in allowing us to construct indexes. The indexes in (1) can be estimated by specifying a demand system and fitting it to the world data, or through a finite, nonparametric revealed preference approach pioneered by Afriat (1967), and more recently developed by Dowrick and Quiggin (1994).

The assumption of identical homothetic tastes is absurd; Afriat (1972) notes that to make such an assumption leaves "the significance of such calculations quite obscure, even as to the locus of injustice." Although identical non-homothetic tastes are usually assumed in trade theory, we would argue this is no more plausible as a description of the world. It is hard to believe that, given identical incomes and prices, demand patterns of Bolivia, Ethiopia, Indonesia, Poland and the US would be identical; perhaps if identical

incomes and prices were maintained for hundreds of years, tastes would adapt, but that is hardly relevant for the present. If tastes are identical but non-homothetic, then the COLI-based system of PPPs depends on a reference utility level—effectively real income—and this modification is not insubstantial. For example, it is not clear that it makes sense to compare Mexican and Peruvian prices as if both were as rich as Japan, or even at some mean level of world income. Neary's (2004) GAIA system of PPPs, which is the leading example of the approach, is constructed on this basis, with explicit assumptions about tastes, and a common system of demand functions estimated worldwide. Identical tastes can be further modified to include "taste-shifters," such as temperature or rainfall, but this just extends the number of reference characteristics that need to be fixed and further stretches the credibility of the numbers. It seems odd, to say the least, to compare the relative costs of living in, say, Congo and Ghana under the supposition that both have the Russian climate. Indeed, many of us would argue that price indexes are not always the same thing as the cost-of-living indexes. If all prices were identical in Moscow and in Ougadougou, it seems reasonable to assert that the *price level* is the same in both, even if the *cost-of-living* were higher in the colder, northern city. In the rest of this paper, we shall follow the national income accountants' practice of thinking in terms of price indexes, not cost-of-living indexes.

One element of cost-of-living theory remains useful in the international context; this is Diewert's (1976) concept of a superlative price index. Superlative indexes, such as the Fisher ideal index, or the Törnqvist index, are defined as indexes that are consistent with preferences that are flexible enough to provide a second-order approximation to arbitrary preferences. Such indexes, unlike the Paasche and Laspeyres, construct indexes using

weights that depend on both the reference and comparison situations, and automatically satisfy the reversal property, that the price level in *d* based on *c* is the reciprocal of the price level of *c* based on *d*. Diewert (2001) also shows that superlative indexes can be regarded as "symmetric means" of the two different indexes that we would otherwise have, just as the Fisher index is a symmetric average (here geometric mean) of the Laspeyres for *d* based on *c*, and the reciprocal of the Laspeyres for *c* based on *d*, which is identical to the Paasche for *d* based on *c*. Since we shall use the Fisher as one of our running examples, these relations are worth recording, and also allow us to establish some notation. With *M* countries, labeled *c*, *d*, etc, and *N* goods, labeled *i*, *j*, *k*, etc, the Laspeyres and Paasche for *d* relative to *c* are

$$P_L^{cd} = \sum_{i=1}^{N} s_i^c \frac{p_i^d}{p_i^c}; \quad P_P^{cd} = \left(\sum_{i=1}^{N} s_i^d \frac{p_i^c}{p_i^d} \right)^{-1} = P_L^{dc}; \quad P_F^{cd} = \sqrt{P_L^{cd} P_P^{cd}} = \sqrt{P_L^{cd} / P_L^{dc}} \qquad (2)$$

where s_i^c is the share of expenditure devoted to good *i* in country *c*. Note that, in ICP practice, these shares come from the National Accounts of each country, while the prices are collected by the ICP itself.

We need one more step to convert the bilateral indexes into multilateral indexes, which we discuss below. But that final step is less important than understanding the implications of (2), and the real conceptual problems that are involved in using the Fisher in particular, rather than some other superlative index. When relative prices differ across countries, different index number formulas will give different answers, and if we cannot rely on identical tastes and cost-of-living indexes, we have a wide margin of choice, which can be thought of as a large margin of statistical uncertainty. One standard way of assessing the size of that margin is to look at the ratio of the Laspeyres price index to the

Paasche price index, the ratio that is being "resolved" by using the geometric mean that gives the Fisher index. Table 1 shows the Paasche and Laspeyres index between the US and selected other countries, in the left panel, and between Nigeria and selected other countries, in the right panel. These are prices indexes for GDP excluding the balance of foreign trade calculated from the 2005 ICP data. Not surprisingly, Canada and Western European economies are closest to the US, and the Laspeyres indexes for prices in those countries relative to the US are only a few points greater than the Paasche indexes. Similarly, Nigeria is "close" to a number of its African neighbors, and more surprisingly, three countries in Eastern Europe, Latvia, Albania, and Estonia also show low spreads. But it is among the countries in the next panel that we see the problem. The US-based Laspeyres indexes for Tajikistan and Kyrgyzstan are 9.6 and 5.1 times the corresponding Paasche index; and although these two countries—especially Tajikistan—are outliers—other countries in Africa and the CIS have ratios more than 2. With spreads that are this large, the choice of price index can make a very large difference in bilateral comparisons, and in some cases, we might reasonably doubt whether the data support any such calculation. We return to these issues below.

Equations (2) give us a set of bilateral superlative indexes which need to be made into multilateral indexes. If we denote by A the M by M matrix of logarithms of the Fisher indexes, we have guaranteed—by the superlative property, that A is skew-symmetric, and it has zeros along the diagonal. However, in general it will not be transitive, in the sense that, for all c, d and e,

$$a_{ce} + a_{ed} = a_{cd} \tag{3}$$

which is what we need. It is straightforward to show that transitivity holds if, and only if, there are M numbers, interpretable as the logarithms of the PPPs, such that, for all c and d,

$$a_{cd} = b_d - b_c \tag{4}$$

There is no principled (backed by economic theory) way of enforcing (4). Gini (1924) suggested choosing b to fit the calculated A by minimization of least squares distance, a suggestion repeated later by Eltetö and Köves (1964), and Szulc (1964), after whom this EKS method is (somewhat unjustly) named. Given that we need to choose $b_1 = 0$ for the base country, the solution is readily shown to be, in terms of the original price indexes,

$$P_F^c = \left(\prod_{j=1}^{M} P_F^{1j} P_F^{jc} \right)^{\frac{1}{M}} \tag{5}$$

so that the EKS-Fisher multilateral index, one for each country, is derived by taking the the geometric mean over all of the possible M "indirect" Fisher indexes from the base country to the country in question.

Many PPPs, such as the Eurostat-OECD and World Bank PPPs, are calculated using some version of the EKS method outlined above. In practice, it is surely impossible to do without the transitivity assumption; we cannot feasibly work with a *matrix* of price indexes. Yet transitivity comes at a serious price, specifically that the price index for any pair of countries depends on prices and budget shares in third countries, a violation of "the independence of irrelevant country" property. Indeed, Van Veeelen (2002) has shown that, given other mild conditions, transitivity and the irrelevance property are mutually inconsistent. As has been known at least since Fisher, price indexes cannot satisfy all of the properties that our price-based intuition suggests for them: price *indexes*

are not *prices*. One possible source of comfort is an observation based on experience, but without theoretical foundation, which is that the matrix of bilateral superlative indexes, such as the Fisher indexes with which we began, are usually close to being transitive without further adjustment, so that the EKS step has little effect on the calculations. By the same token, comparisons between pairs of countries using EKS price indexes are not very sensitive to prices or budget shares in third countries. We also have a nice compromise between statistical practice and cost-of-living theory, in that the Fisher indexes are superlative indexes with a COLI interpretation if we are prepared to make the assumption of identical tastes, at least for some countries.

These indexes have (at least) one compensating drawback, which is responsible for the fact that they are not used in the Penn World Table. The Penn World Table is a set of national accounts, with consumption, investment, GDP and so on, and these national accounts satisfy the standard national accounting identities, for example that consumption plus investment plus government plus exports minus imports is equal to GDP. When each component is converted to PPP by using an EKS index for each, and expenditures converted to international currency by division, these identities no longer hold. Components of aggregates in international prices do not sum to their aggregates in international prices. For some purposes, for example for the World Bank's poverty work which depends only on the PPP for household consumption, this is of no consequence. And Eurostat estimates appear to be widely used, at least within government agencies, without satisfying this requirement, though Eurostat also produces alternative estimates with a lag. But economists studying the structure and dynamics of macro economies

might find the violation of identities to be disconcerting, and the PWT is constructed along principles that preserve them.

The aggregation formula used by PWT was proposed by Geary (1958), and is usually referred to as the Geary-Khamis (GK) method, Khamis (1972). In the GK system, the prices in each country are compared with those of an imaginary composite country, itself constructed from averaging the countries in the system. In the case of two countries, the GK PPP index is computed as a Paasche index that compares domestic prices with "world" prices, which are the prices of the composite so that, for $c = 1, 2$,

$$P_{GK}^{c} = \frac{\sum_{n=1}^{N} p_n^c q_n^c}{\sum_{n=1}^{N} \pi_n q_n^c} \tag{6}$$

where π_n is the world price of good n, which is itself defined as the quantity weighted average of the prices of good n in each country, expressed in the global currency:

$$\pi_n = \sum_{c=1}^{M} \frac{p_n^c}{P_{GK}^c} \frac{q_n^c}{\sum_{d=1}^{M} q_n^d} \tag{7}$$

Equations (6) and (7) must be solved simultaneously, which can be done iteratively, or as shown by Diewert (1999) as the solution to an eigenvalue problem.

The advantage of this Geary-Khamis system of PPPs is that it preserves aggregation; because there exists a world price for each good, each item of GDP is re-priced at the world price, and added up to give re-priced subgroups or totals. The GK system also has a number of disadvantages that need to be balanced against this. Unlike the indexes underlying the EKS approach, it is not superlative. In consequence, if the two countries had the same homothetic tastes, the Geary-Khamis index would not be a second-order

approximation to the "true" cost-of-living index. If this were the main concern, Geary-Khamis could be replaced by Neary's (2004) GAIA system which is a consumer-theory consistent version of Geary-Khamis. If we do not want to assume identical tastes, nor use the cost-of-living framework that would be arguably appropriate if tastes were identical, these objections to Geary-Khamis are not decisive.

To see the central problem, note that the quantity weighting of prices in (7) means that the country with the larger physical volume of consumption of a good gets greater weight in the construction of the composite world prices. If, for example, we used Geary-Khamis to compute a PPP for Bangladesh relative to the US, the world prices would be close to those of the US. In the Penn World Table as a whole, Nuxoll (1994) has argued that the composite world prices are those that would characterize a middle-income country such as Italy or Hungary. The use of such prices has the effect of overstating the level of consumption in poor countries. For example, many services—haircuts, domestic service, restaurant meals—are cheap in poor countries because people are poor, because such services cannot be traded, and because labor is not free to move around the world. If we use (say) Italian prices to value (for example) Indian consumption, these components of consumption will be valued very highly, and will inflate the value of Indian consumption at international prices. This is called the Gershenkron (1947) effect, the overvaluation of one country's consumption when evaluated at another country's prices. Put another way, it is the understatement of the price level in one country relative to another that comes from using a Paasche index; in (6) the domestic value of consumption in the numerator is divided by the inflated value of consumption at world prices in the denominator. Of course, the understatement of the Paasche index is an understatement

relative to a true cost-of-living index, which is itself not well-defined without common tastes. But if we compare the Paasche in (6) with the superlative indexes presented above, the latter always averages weights from *both* countries, whereas the GK indexes use only *domestic* weights, and this is what generates the Gershenkron effect. Goods from rich countries are often rare and expensive in poor countries, if they exist at all, and goods which are rare in countries with no taste for them—alcohol in Muslim countries or English sausages in the US—can be very expensive when they are found at all. If alcohol has a small share in Bangladesh, but a high relative price, weighting that price relative by Bangladeshi budget shares will understate Bangladeshi prices, while weighting it by OECD shares will greatly overstate them. The superlative indexes, which combine the weights, make a compromise that is arguably the best that can be done in the circumstances.

Compared with the superlative indexes, GK indexes will understate PPPs in poor countries relative to rich ones, and overstate their living standards. They make the world look too equal, and understate poverty in the poorest countries. For analyses of the world distribution of income or of world poverty, this would militate against using GK indices and in favor of EKS type methods. For work on growth on or other macroeconomic questions, the additivity properties of GK are likely to be more important, especially when the analysis is dominated by rich countries. When such analyses involve poor countries in a substantial way, it should be borne in mind that the international prices that are used to value their goods and services are biased towards rich country prices, with the risks that this entails, for example in overvaluing cheap goods and services that have relatively little domestic value.

The EKS and GK are not the only methods for generating multilateral price index numbers, but they dominate in the regularly used data bases, and so we do not discuss the other formulas here.

Table 2 illustrates the various indexes for the same selected set of countries as in Table 1. Column 1 repeats the Laspeyres–Paasche spread as a reminder of where we would expect the different multilateral indexes to differ; the EKS indexes are modified Fisher indexes, and the GK modified Paasche indexes and will inherit at least some of their properties. We also show the bilateral Fisher index with the US as base, and then our own calculations of the EKS and GK indexes using 128 basic headings of GDP, excluding only the trade balance. The final column is the PPP index from the World Development Indicators 2008, which we shall discuss below.

There are several points to notice. First, the bilateral Fisher indexes and the EKS indexes are not far apart, so that the EKS modification to the bilateral indexes is not having very much effect. If this is true more generally—as we suspect is the case—the EKS indexes can be thought of much as we would think of Fisher indexes. The GK indexes are further away from both Fisher and EKS indexes than they are from one another. Even so, the differences are small for similar countries where the Laspeyres–Paasche spread is small, but can become substantial where the spreads are large, or in other cases where the data are weak, such as Zimbabwe or Tajikistan. In the worst case—Tajikistan—the GK index is only 55 percent of the EKS. The ratio of EKS to GK is 1.056 in Nigeria, which is probably the worst of the large countries.

Whether these differences are significant depends on the use to which the data are to be put, on which we shall have more to say below. One illustration comes from looking at

world inequality, or at least the (dominant) between-country component of world inequality. We compute this from the local currency value of per capita GDP, and deflate by the EKS and GK indexes. As predicted, inequality is smaller using the GK PPP, although by very little; the population weighted gini coefficient for per capita GDP is 0.533 for EKS, compared with 0.527 using GK.

3. Operational issues with major implications for use of the data

3.1 Prices of items, prices of basic headings, and quality

The ICP collects and constructs the prices that go into the formulas, and some understanding of this process is necessary. As is the case for domestic prices, such as the CPI in the US, there are two levels. At the bottom level, ICP investigators price thousands of items using detailed lists of closely comparable products. In the 2005 ICP, this was done separately for regions of the world, each armed with its own list, so as to avoid problems of looking for European items in Africa or vice versa. These item prices are then aggregated up to 129 basic headings, each of which is a component of GDP approached from the expenditure side; 106 of these basic heads are items of consumers' expenditure. Below the basic heading level, there are no expenditure or quantity data, so the detailed prices have to somehow be aggregated up to prices for basic headings without relying on weights, and this process turns out to have important consequences for the end result. Once we get to basic headings, we have expenditures from each country's national accounts, and these expenditures can be used with the prices for the basic headings to fill out the formulas in Section 2, which use expenditures, expenditure shares, or quantities—expenditures divided by prices—to weight the prices. The prices—or

better price indexes—for the basic heads are often referred to as "parities," and are usefully thought of as commodity-specific PPPs. For example, "rice" is one basic heading in the consumption account. Some illustrative parities for rice from the 2005 round are 4,304 Vietnamese dongs per dollar, 0.65 British pounds per dollar, or 44.6 Kenyan shillings per dollar. If rice were the only component of consumption (or GDP), these would be the PPP exchange rates for those countries relative to the US; in fact, the actual consumption (GDP) PPPs for those countries are 5,920 (4,713) Viet Nam, 0.66 (0.65) UK, and 32.7 (29.5) for Kenya. Clearly, knowledge of the price of one good, or at least one group of goods, takes us a long way, which is why the Economist's Big Mac index is useful. Of course, the ICP tries to do better than this by covering all the expenditures in GDP.

In Section 2, we discussed the aggregation of basic headings into overall indexes for GDP and its components, but just as important—and considerably more difficult—is the stage below the basic heads where individual goods are priced and aggregated up to the basic head level. This is the area where comparability is most difficult, where the ICP has to face the issue that not all goods are available everywhere, and where the vexed issue of quality comparability come into sharpest focus. One persistent criticism of ICP rounds prior to 2005 has been that the quality of items priced has not been strictly enough controlled, so that lower quality items in poor countries were often matched to higher quality items in rich countries, leading to an understatement of price levels in poor countries and to an overstatement of their output and income levels. This is more pervasive than just the concern that brain-surgery in Nairobi is unlikely to be identical to brain-surgery in Geneva, because for many goods the outlets sampled in poor countries

may be closer to "dollar-stores" than to the typical outlet in the US. In consequence, successive rounds of the ICP have developed ever more precise descriptions of the goods to be priced, leading in 2005 to a formal set of descriptions known as "structured product descriptions."

Each region of the ICP has a list of goods, with their structured product description, and there are typically many of these goods within each basic heading. Even within a region, not all countries will be able to price all items. Clearly, this selection poses potential problems, though there are problems in the other direction if countries price goods that are rare and expensive. These difficulties are further exacerbated by the lack of expenditure information below the basic heading level, whose presence would be useful for weighting and could act as a guide to the importance of price anomalies.

To deal with these issues, Eurostat-OECD and the CIS developed procedures in which national statisticians judgmentally marked goods as either representative or not representative in each country. When the price quotes for the goods were combined to construct the bilateral price indexes for basic headings between countries that are the first stage of the EKS construction, care is taken to avoid comparing representative goods in one country with unrepresentative goods in another, see World Bank (2008, pp 157–8 for details.) Although this procedure is *ad hoc*, so will be any other way of averaging prices within basic heads, and it provides some protection against the use of artificially high prices of items ("exotica") that are only rarely found in some countries. In the 2005 round of the ICP, it was planned to use representativity for all of the regions, but the attempt failed. Taking this together with the stricter control on the specification of matching goods, the direction of the quality bias is no longer clear, if indeed it ever was.

Many of the qualities available in poorer countries are not available in higher income countries, while more of the qualities available of richer countries can also be found in poorer countries. Although prices are sometimes collected in outlets specially selected for the ICP, the outlets used for the local consumer price index are generally used, at least as a starting point. In addition, many of the higher quality items are international brands while regional or brand-less products are more important for lower quality items. The consequence is that higher quality items tend to dominate the actual list of items compared in the ICP if only because the initial specifications were drawn heavily from the EU-OECD region. Many of these items will not be in the CPIs of poorer countries and will often not be available in the outlets normally sampled for their CPIs, an effect that is likely to be exaggerated because the CPI outlet and item samples in many developing countries are out of date, often by several decades. The consequence is that prices for the ICP were often collected in higher-end outlets, which has the effect of raising price levels of poorer countries. This was made more likely in 2005 than previously because of the much closer review of prices across countries so that, for example, international brands were priced in (say) China, because they were available, even if mainly in high-end outlets. To the extent this happened, it would have the effect of raising parities in poorer countries, making them appear to have less income and output than in fact they do.

Paradoxically then, continuing attempts to control quality by more careful matching may possibly have gone too far, and the 2005 ICP may be overstating prices and understating income levels in the poorer countries. These quality issues are likely to have contributed in some part to the apparent reductions in the size of the Indian and Chinese

economies, though other factors are almost certainly more important. In future, it is possible that greater use of hedonic techniques—for which the structured product descriptions are well suited, see Kolkowski, Moulton, and Zieschang (1999)—may help in matching qualities without so much emphasis on matching narrowly defined commodities. But it is also worth referring back to our opening example of comparing the *teff*-eater with the rice-eater, for which precise specification will only make the comparison impossible. One possible solution is to think about the characteristics that such goods provide, in terms of proteins, calories, and micronutrients per unit of currency, and constructing PPPs for those characteristics, not qualities of goods, see for example Deaton, Alatas, and Friedman (2004) for an attempt along these lines..

3.2 China, India and other large countries

Until 2005, China had never participated in an ICP round, so that in all versions of the PWT to date, the Chinese numbers were estimated using partial information and shortcut methods. India, which also participated in 2005, had not previously done so since 1985, and its prices since then have been estimated by a mixture of updating and regression methods of the kind described in Section 4 below. The estimates for both, whether in the PWT, or other PPP databases, are therefore relatively uncertain compared with recently benchmarked countries. The Chinese and Indian data from the 2005 round, which will be incorporated into version 7.0 of the Penn World Table, are therefore of more than usual interest. Discussion of China and India also brings up an important general issue, which is differences in prices across space, both between different cities, and between urban and rural sectors within a country. In principle, prices for the ICP are national average prices,

but in practice, rural prices are not always collected, something that is a more serious issue in large economies than in small ones.

We start with China, where the 2005 price levels are much higher than previously projected, so that there is an immediate question whether prices were overstated through some combination of choice of brand, outlet or location. Price collection by China in 2005 took place in 11 cities and in their immediately surrounding areas which are mostly urban, but with some rural characteristics. Apart from these, no rural prices were collected. Price differences between the cities or between the cities and their surrounding rural areas were not reported, but they apparently were not large. Because the Chinese expenditure data refer to the whole country, the Asian Development Bank (ADB), which was the regional authority that handled the Chinese data, adjusted the urban prices to an all-China basis to match what was done in fully participating countries, Dikhanov (2007). These adjusted prices are weighted averages of the data from the 11 cities, which at best can correct for lack of national representativity of the cities, but does not deal with any difference between urban (or urban plus periurban) and rural prices.

That urban prices are substantially higher than rural prices has been documented by Brandt and Holz (2006, Table 7), who estimate that the overall ratio of urban to rural prices was 1.33 in 2000, and by Chen and Ravallion (2008) who estimate an urban to rural price ratio of 1.37 for 2002. Neither of these estimates fully captures differential housing costs, so that these estimates are too low. Chen and Ravallion also note that the Chinese National Bureau of Statistics chose the 11 cities because they were most likely to have outlets carrying the types of products and brands in the ICP specifications, and those prices are likely to be unrepresentatively high. Taking all of this together, we hazard a

guess that the urban prices as measured are perhaps 40 percent higher than rural prices. Around 40 percent of aggregate consumption takes place in rural areas, so that the PPP deflator for consumption is overstated by perhaps 16 percent, and for GDP by a little less than 10 percent.

Do similar spatial differences exist in other large countries? Aten (2006) reports that for the 38 urban centers used by the US for the CPI the differences between small southern urban areas and San Francisco are large, 80 versus 130 percent of the US average in 2003. From more than a million collected prices, Aten is able to obtain about 25,000 annual average price observations for 256 entry-level items collected by the BLS and uses these to estimate price level differences over all of consumption. This is a rich data set that has now been updated to include 2004 and 2005 with similar findings, so that we can be fairly certain that the range across US urban areas is around 60 percent, suggesting that Brandt and Holz's estimates for China, where spatial price differences are almost certainly larger, are not unreasonably large. Aten also finds that the gradient of prices from low to high is not large for goods, but it is much steeper for services, a common finding of previous rounds of the ICP across countries. Unfortunately, it is service items like housing, medical, and personal services that have not been surveyed or measured very well in the ICP, nor in the expenditure surveys that underlie the Brandt and Holz study.

There is a clear problem for the ICP in comparing large versus small countries. If all countries had the same ratio of urban to rural prices, the PPPs based on urban samples of prices would be the same as those for national prices. For Belize, Bermuda, Hong Kong, Luxembourg or Singapore, the frame of outlets for the CPI covers the country and

provides a good basis for the ICP. But this is much less true for large countries. Not only are large parts of these countries not covered in the CPI, but the sample of outlets is typically not well suited to the ICP's lists of items—the problem of comparable goods once again. In consequence the degree to which large developing countries rely mainly on urban prices varies greatly across countries. India has a long tradition of collecting rural and small center prices, while Brazil, Indonesia, Pakistan and Thailand have typically collected urban prices. Even in India, the usual outlets for collection of rural prices would not typically include outlets in smaller regional centers where some of the ICP items are more likely to be available.

Until there has been more analysis of the results not much more can be said about how we should interpret price levels between small and large developing countries as reported in the 2005 ICP.

3.3 Comparison resistant items: housing and government services

The ICP uses the term "comparison resistant" to refer to goods and services for which it is difficult or impossible to observe market prices that can be compared across countries. Examples are housing rental, government services, health, and education. Measurement of these is problematic for the national accounts within countries, but becomes still more hazardous in international comparisons. We focus on rental and government services, but will have something to say about heath and education too. Many macroeconomists using these data may be interested only in aggregates such as national income or consumption, and not in these items for their own sake. But the measurement problems for these items are large enough to have major effects on the larger aggregates; indeed the largest single

factor responsible for the decline in the relative size of the Indian and Chinese economies is a change in the treatment of government services.

How does one compare the output of civil servants and health and education workers across countries? Within countries over time, national accounts deflate nominal salaries to a quantity basis by making assumptions about changes in the productivity of the relevant workers. Similarly, in the ICP, assumptions about differential productivity across space fill in for the missing prices. In past ICP rounds, volumes were derived by dividing compensation by a PPP derived from a detailed comparison of salaries for occupations typical of government staff expenditures. Such an assumption implies that productivity in the provision of these services is identical across countries in a given occupation, which is unlikely given very different amounts of accompanying capital per worker across countries. Further, there is little inducement to organize the work environment to improve productivity of employees in administrative, health and education services in very low-wage economies.

In the 2005 benchmark, the range of countries was much greater than in previous rounds, so that the consequences of the equal-productivity assumption loomed much larger. In Asia for example, salaries for government health workers measured at market exchange rates are 120 times larger in Hong Kong than in Laos; similar differences exist between Yemen and Kuwait in the Western Asia region. If we assume that productivity is the same, per capita volumes of these comparative-resistant services in Yemen or Laos greatly exceed those of its richer neighbors, an improbable finding. Such adjustments have been considered earlier by the OECD and the ICP, but the 2005 comparisons in the Asia-Pacific, Africa, and Western Asia regions are the first actual cases where the equal

productivity assumption has been significantly modified. Productivity differences in OECD-Eurostat, CIS, and South America were not dealt with, either because they were thought to be too small to warrant it or because there was no agreement on how the correction should be made.

Asia, West Asia and Africa made these adjustments based on estimates of capital per worker in each economy as a whole. In Asia, for example, this has meant that the volume of GDP of China and India relative to Hong Kong or Singapore is lower than in previous ICP rounds. Heston (2008) estimates that the effect of the adjustment was to reduce Asian GDP by 12 percent relative to the OECD countries, and Indian and Chinese GDP by much more, 15.6 percent and 14.5 percent, or nearly 40 percent of the total downward revisions; this calculation is based on an aggregation over all countries simultaneously, and may not hold for the regional aggregation actually used in the 2005 ICP, see the next subsection. There is therefore a comparability issue across regions in 2005 because Europe, the OECD, the CIS, and South America made no such adjustments. Further, because capital per worker data were not available for many countries, it was often necessary to apply the same adjustment factor to low- income countries that were at different stages of development. The actual procedure used is described in the final Reports of Asian region and the ICP as a whole, Asian Development Bank (2007) and World Bank (2008a).

The productivity adjustment is clearly in the right direction relative to earlier benchmarks which attributed too large a volume of such services to poorer countries and biased upwards their PPP converted GDPs. However, the particular procedure was based upon limited information applied uniformly over groups of countries within each region,

so there is an unknown, but likely significant error associated with the actual adjustments, even for countries within the same region. Further the adjustments in Africa and West Asia were each calibrated differently than for Asian countries. What does this mean for comparing the 2005 results to previous benchmarks? In previous benchmarks, the volume of administrative, health and education services for very low wage countries in Africa, Asia, and Western Asia would have been substantially lowered if the 2005 procedure had been adopted in those earlier years. The new procedures have thus artificially widened the distribution of income between countries.

What is the consequence for the 2005 comparison of the mixed application of an adjustment for productivity in some regions and not in others? Certainly Asian GDP (excluding Korea and Japan) was reduced compared to the OECD countries (here including Korea and Japan) as a consequence of the productivity adjustment. This means that comparisons of Asian countries with peer countries like Brazil, Mexico and many eastern European countries where the productivity adjustment was not carried out, would also be affected. This is not an argument against a productivity adjustment, though the actual implementation was perhaps insufficiently tailored to country specifics. But a knowledge of the actual adjustment helps us better understand why the position of China and India has changed so much in the recent round. And it certainly points to the need to gain agreement on a standard method of treating un-priced services for future rounds.

Another important comparison-resistant component of GDP is the rental of housing, including both actual rents and the imputed rents of owner-occupiers. The share of housing rent in GDP is about 10 percent in the US, 9 percent in the UK, but only 4.7 percent in India, 2.2 percent in Nigeria, and an incredible 0.5 percent in Ghana; some of

these differences are likely to reflect difficulties in measuring rent in the national accounts, or failure to make any imputation at all for owner-occupier rents. .

Prior to 2005, ICP comparisons were based on surveys of rents, which allowed market rent comparisons for various size and amenity groups of housing and, assuming rental equivalence, for owner occupied housing. The EU and OECD countries used a similar survey approach until their expanding memberships included countries that were not suitable for surveys of market rents. A new member country might have a small expatriate community that paid market rents, and if other rentals existed, they were subsidized. The approach of the EU for such countries was to make direct comparisons of quality-adjusted volumes of housing, the "quantity" as opposed to the "survey" approach, and to find a link member country or countries, initially Austria, that would both survey rents and provide quantity information on their housing stock. Of course, other countries had faced similar problems earlier, including those with rural housing stocks that are rarely if ever rented and had adopted a range of methods.

For the 2005 comparison the plan was to use both a quantity and survey approach or some combination in other regions. In practice the quantity approach was used in South America, and a combination in Western Asia. In Asia, however, neither approach appeared feasible for all countries so measurement was effectively abandoned in favor of the assumption that the per capita volume of housing services for each country was the same proportion of the world average volume of housing services as was the remainder of actual household consumption. The same approach was adopted in Africa. One consequence is that it is not meaningful to compare housing volumes in any country in Asia and Africa with a country in the other regions. Another is the bizarre consequences

in countries, like Ghana, whose national accounts show little expenditure on rents. In these cases, the PPP for the rental category is calculated by dividing a very small number by a relatively large one, so that the parity for this basic heading is wildly out of line with the overall PPP. For Ghana, the parity for the rental heading is 178, compared with an overall PPP of 3721, less than 5 percent. This is an extreme case, but there are others that are almost as extreme; Malawi also has a ratio of 5 percent, and Chad a ratio of 6.9 percent. Although the local (Ghana, Chad, Malawi) expenditure weights attached to these prices are also very small, the same is not true of the countries with which they are compared in calculating the bilateral Fishers that go into the PPPs. For the three countries listed, their overall PPPs are reduced by close to ten percent comparing PPPs with and without the rental category, and the size of their estimated incomes inflated by the same amount. Perhaps this is not very large given the overall quality of African GDP data, but it is worth bearing in mind.

We shall not discuss health and education here, but there is an associated issue that might not be familiar to all economists. The most recent versions of the United Nations' System of National Accounts (SNA), adopt a definition of household consumption, referred to as "actual consumption" which includes services provided for households by government and non-governmental organizations, including health and education. From a national accounts perspective, this makes a certain amount of sense, because it prevents the size of consumption depending on the extent to which health and education are publically or privately provided. However, there are many countries around the world where government-provided health and education is inefficient, sometimes involving mass absenteeism by teachers and health workers, Chaudhury, Hammer, Kremer,

Muralidharan and Rogers (2006), so that such "actual" consumption is anything but actual. To count the salaries of AWOL government employees as "actual" benefits to consumers adds insult to injury. The original, more traditional concept of household consumption, is sometimes presented—though sometimes hard to find—and is typically labeled "individual consumption expenditures by households" or some other term not containing the word "actual."

3.4 Regions

We have at several points noted the regional structure of the 2005 ICP, and indeed data have been collected on a regional basis since 1980. One reason is logistical, and it also allows the ICP to take advantage of data and expertise that already exist, for example in Eurostat. More substantively, PPP comparisons are almost certainly more reliable between countries that have a similar economic structure, whether through tastes or conditions of production, so that regional systems of PPPs are probably more reliable than global systems. Once again, the relative reliability of comparisons between close neighbors is clear from Table 1. As in earlier rounds, the 2005 ICP was decentralized, on this occasion into Africa, Asia-Pacific, Commonwealth of Independent States (CIS), South America, Western-Asia, and Eurostat-OECD. The Eurostat-OECD part of the ICP covered 46 economies in Europe and the non-European OECD, and is best regarded as an independent—although coordinated—effort. In the first stage of the ICP, each region collected prices using its own list—matched to the global list of basic headings—and calculated a regional set of parities for each basic heading, and a regional set of PPPs.

Each region had its own numeraire country, sometimes a real country—Argentina in South America—and sometimes an imaginary composite country.

None of this is controversial although, as we have seen, there were some variations in practice across regions—including the use of the Iklé method (an additive system of PPPs, closely related to Geary-Khamis) in the African region alone. The controversy and difficulties start in deciding how to link the regions. In the previous round (1993), the linking was never satisfactorily completed, severely limiting the use of the data. The idea here was to work with "bridge" countries that appeared in more than one region, just as bridge periods are used to link time-series of prices on different bases. In the 2005 round—although not by original design—the single bridge country was Russia, which was included in both the EU-OECD and CIS regions, and which priced both lists. Russia was then used to link the two regions into a single EU-OECD-CIS region with a single set of basic-heading parities and country PPPs relative to one numeraire. In 1993, one major problem was the sensitivity of the linking to the economic structure of the bridge country, so that in cases where there was more than one possible bridge country, the results were quite sensitive to which was chosen, and further sensitive to the level (basic heads or whole country) at which the linking was done.

In the 2005 round, bridging was replaced by an upper-level "ring," consisting of two or more countries in each region, each of whom agreed to price a distinct "ring" list. As suggested by Diewert (2008), these prices were then used to estimate a set of region multipliers, using a regression of the form

$$\ln p_{ig}^{cr} = \alpha_c + \beta_r + \gamma_i + \varepsilon_{ig}^{cr} \tag{8}$$

where p_{ig}^{cr} is the price of good *i* within basic heading *g* in country *c* in region *r*. The estimates β_r are the regional effects that allow us to link the regions into a global system; if p_i^c is the parity for basic-heading *i* in country *c* relative to the numeraire in region *r*, its global parity will be $p_i^c \exp(\beta_r)$. (Although the CIS was supposed to be included in the ring, it declined to do so, which is why the CIS was combined with EU-OECD using Russia as a bridge.)

When the Penn World Table 7.0 is available, it will contain a set of GK accounts based on these global parities. However, the currently available numbers for 2005 that appear in the ICP Final Report, as well as the time-series in the 2008 WDI, are not constructed in this way. Starting in the 1980 round with Eurostat, whose freedom of action is heavily restricted by EU regulation, all regions from 1985 onward made their participation in the ICP dependent on the condition that their own regional PPPs were respected in the final calculations, a restriction known as "fixity." As a result, the PPPs that appear in the World Bank data for EU-OECD are identical to those calculated by Eurostat up to a scale factor. In the 2005 round, equation (8) was estimated in a way that respected fixity, by taking the previously determined α_c, which are the logarithms of the within-region PPPs, over to the left-hand side of the equation before estimating the regional factors. This means that the parities for the basic headings, which are available to researchers for further analysis, have one aspect of fixity built into them through this treatment of equation (8). Similarly, the new versions of the PWT must use those parities for the basic headings, even though it is under no obligation to respect regional fixity in the future, and it has not done so in the past. But at the very least, the ignoring of fixity is one of the reasons why the PWT differs from other PPP data sets.

We are not aware of an analysis of theoretical advantages and disadvantages of imposing fixity, which comes as a political, not conceptual, constraint. In the limit, if every country demanded fixity of its price level relative to all other countries, transitivity would be impossible, and we would not have an international system of accounts. Actual fixity is clearly less severe, but it also places restrictions on the way that transitivity is imposed, and so will cause the final global system of PPPs to be further away from the matrix of pairwise superlative Fisher indexes that would be the basis for an unrestricted calculation for all countries simultaneously. On the other hand, it is also clear that the Fisher indexes between two countries in different regions, between the US and Tajikistan, or between Nigeria and Japan, say, are worth less than those between the US and Canada, or Nigeria and South Africa, see again Table 1, and fixity recognizes this fact, albeit in a crude way. The PPPs from the ICP, which incorporate the fixity constraints, are listed in the final column of Table 2. For several of countries, these estimates are outliers relative to the other indexes shown. This is true, not only where it is to be expected, in the countries with very large Laspeyres-Paasche spreads, but also in some of the important countries in the bottom panel. For example, the PPP for China in the ICP is more than ten percent higher than the EKS or bilateral Fisher index, while that for India is 7.5 percent higher. These results are somewhat puzzling given the EKS basis of the Asian numbers, more so than the also large difference for Nigeria, where the ICP does not use the EKS method. We suspect that the interaction of the productivity adjustment and the regional aggregation may be largely responsible for the differences. It is perhaps also worth noting that the use of the PPPs from the ICP also has a mild positive effect on measures of inequality between countries; the population weighted gini coefficient of GDP per head,

which was 0.533 and 0.527 for the EKS and GK becomes 0.580 using the ICP official data.

4. Filling in over space and over time

Since the ICP was begun in 1968, there has been an interest in covering non-benchmark countries, and PWT was launched to serve that interest. In addition the international agencies wish to cover as many of their member countries as possible. A brief discussion of filling in non-benchmark estimates is given in section 4.1. This involves both spatial and temporal extrapolation. Section 4.2 is about the extent to which successive revisions of the PWT are compatible, and reports important lessons from recent research undertaken at the IMF and elsewhere. In these subsections, we give most attention to the PWT, since it has been most heavily used by in economic research, and because it is the only one of the databases that provides a long time series of disaggregated national accounts for the world as a whole. (Similar data for Europe and the OECD, stretching back to 1980, and including some higher frequency data—quarterly and monthly—come from the Eurostat-OECD PPP program)

4.1 Filling in over space and over time

When non-benchmark estimates were launched in the mid 1970s the number of benchmark countries was only 16. In 2005 the task is much less daunting because there are 146 benchmark countries, so the number of non-bench mark countries is about 40, many of which are in the Caribbean which was the only significant country grouping that did not participate. Different databases use different imputation rules for the countries

without data, although all rely on the strong positive relationship between national income and the price level of GDP, defined as the ratio of the PPP to the market exchange rate. Poorer countries are relatively richer in PPP terms, proximately because non-tradable goods are cheap where wages are cheap, and more fundamentally, because of the Balassa-Samuelson conjecture, that rich countries are relatively more productive in the traded-goods sector. Figure 1 shows this relationship for the 2005 round, plotting the logarithm of the price level of GDP, with the US as 0, against the logarithm of GDP per capita expressed in market exchange rates; the heteroskedasticity in the Figure may reflect data quality as much as failure of Balassa-Samuelson at low incomes. Each country is plotted with a circle whose diameter is proportional to population size; the gross outlier here is Zimbabwe. The World Bank runs this regression in levels, somewhat oddly given the graph, with the price of gross national income as the dependent variable, and with high school enrolment rates as an additional explanatory variable. (We understand that the Bank plans to improve this procedure in future.) The PWT uses the log of the price level of domestic absorption as the dependent variable and adds, not education, but direct (although admittedly imperfect) information on prices taken from foreign-posting cost of living adjustment indexes from the International Civil Service Commission (ICSC), as well as from the US and Canadian foreign offices. PWT also uses an openness to trade variable, a variable that captures involvement in international financial flows, and regional dummies for Africa and the OECD. Conditional on these variables, education plays no significant role.

The main reason that PWT uses domestic absorption to impute PPPs is that handling of trade is weak in both the benchmark comparisons as well as in extrapolations over

time. Turning to the latter point first, the WDI relies on extrapolations of the price of GDP, and this is one reason that their estimates for 2005, which relied on extrapolation from 1993 benchmarks, were often markedly different from the new benchmark estimates. For example, if export volumes are constant but their prices fall, as in the case of micro-chips for Singapore, GDP growth will overstate the ability of Singapore to convert current production into current domestic expenditures in 2005. Other changes in the terms of trade will similarly drive a positive or negative wedge between extrapolations and current price PPP conversions.

The net foreign balance in PWT is converted at the PPP for domestic absorption which, as argued in Feenstra, Heston, Timmer and Deng (2009) means that the PWT should be regarded as providing estimates of income, not of output. When the net foreign balance is converted at exchange rates, as in the 2005 ICP, countries are treated differently depending on whether the foreign balance is positive or negative and their price level is greater or less than 1. For example, the PPP for both China and India is less than their exchange rate. China's surplus is therefore converted at a value less than its command over goods in the Chinese market, while India's deficit is converted to be a smaller reduction than it would be at domestic currency. If we want to obtain real output, we must convert exports and imports at their own PPPs, which was the exercise that was implemented using unit values for 1996 by Feenstra, Heston, Timmer, and Deng. They find significant differences between output and income suggesting this is a fruitful line of research to pursue, notwithstanding the difficulties of disentangling quality from unit value differences. As they note, the distinction between output and expenditure measures

is likely to be important for researchers using growth regressions that are intended to be interpreted as modeling the output side of the economy.

There are a number of other issues in computing changes in GDP between benchmarks in PWT, and some of these are certainly questionable. The calculations start with the current price estimates in PWT. The PPPs for consumption (*C*), investment (*I*) and government expenditure (*G*) for a base year, say 2000, are moved backwards and forwards in time by the ratios of the national deflators of each aggregate relative to the corresponding deflators in the US. Then a new Geary-Khamis aggregation is carried out for say, 1995, that provides a current price estimate of the 1995 PPP for domestic absorption. The net foreign balance in current prices is converted at the PPP for domestic absorption and added in to obtain GDP. This provides a time series in current prices. (This procedure is modified in cases where the benchmark is long out of date, as in India in recent years, in which case the estimate is the above averaged with a regression-based estimate, as if there were no benchmark.)

A number of questions arise about GDP in constant (say 2000) international prices, especially because this is the most commonly employed series by users in their models. PWT has offered two principal constant price measures, a fixed weight and chain weight index. The fixed weight index uses the share of *C*, *I*, and *G* in 2000 as the weights applied to national growth rates for each of these expenditure aggregates. The chain weight index applies the current price weights of the year *t* to the growth of *C*, *I* and *G* between *t* and $t-1$. This provides an estimate of the growth rate of domestic absorption between *t* and $t-1$ to apply to domestic absorption in year *t* in 2000 prices. Many inquiries have been fielded about PWT for the past 25 years but none have questioned the fixed and chain

indexes, which should perhaps best be interpreted as benign neglect rather than critical acceptance. But for a chain index, given the limited national accounts detail available, the above procedure is the only alternative. For a fixed weight index it would be possible to simply use the national growth rate of domestic absorption, and then add the net foreign balance, a series that will be provided in the future.

How should the net foreign balance be converted in constant prices? Up until now in the PWT the growth rates of exports and imports were applied to the 2000 values of exports and imports for each country converted at the PPP for domestic absorption. An alternative treatment was also offered that attempted to take account of the terms of trade. Neither method is satisfactory. National growth rates of exports and imports usually reflect changes in production with fixed weights. Trade is an important area for improvement in both the ICP and PWT.

A final issue worth noting is the treatment of the formerly communist countries of eastern Europe and the former Soviet Union. Until events of 1990-91, PWT followed the prevailing practice of making estimates for “market economies”, and treating “non market economies” separately. Several of the latter, Hungary, Poland, Romania and Yugoslavia had participated in the ICP, but their constant price national accounts were not in a form that could be moved to non-benchmark years. The European non market economies and occasionally Cuba, though never China, had been carrying out purchasing power comparisons since the 1960s under the Council of Mutual Economic Assistance (aka COMECON) making binary comparisons with the Soviet Union. These studies were not officially published until the last comparison for 1990, but they appeared in journals and were the basis of the early treatment in PWT.

The integration of estimates of these countries into PWT began with PWT 5.6. However, there were major obstacles for those COMECON countries that had not previously recast their national accounts from the Material Products System into the SNA form. Further, price indexes over time for many of these countries were particularly difficult because the movement from administered to market prices posed a major data collection problem because there did not exist any framework for collecting prices in the field. Hungary and Poland were well equipped for the transition, but to varying degrees the quality of data is uneven; those former Soviet Union (FSU) and Eastern European countries now members or associates of the EU have the most reliable data for ICP and PWT purposes. An exception would be those areas affected by the break-up of Yugoslavia. The Commonwealth of Independent States (CIS) and its statistical office CISSTAT (the former COMECON secretariat in Moscow) now coordinates PPP and other studies for the nine member countries. Russia, which also participates directly in OECD PPP studies, has a strong statistical apparatus, while the other countries are quite mixed as is illustrated by the case of Tajikistan in Tables 1 and 2. Further the constant price national accounts series for the COMECON countries are all much more reliable after the mid to late 1990s than earlier, with Hungary and Poland as (positive) exceptions.

4.2 Is PWT consistent across versions?

In his *Principles of Economics*, Marshall began with a variation of the Latin proverb, *natura non facit saltum,* there are no jumps in nature. Reviews of studies based on different versions of PWT raise questions about whether some of the jumps in PWT are real or rather in the nature of PWT's construction. As PPP estimates have evolved over

the years the underlying data-base of PWT has also been revised. New benchmark estimates bring in additional countries and revised estimates for multiple benchmark countries. As we have seen from the discussion of the 2005 round of the ICP, substantial methodological changes continue to be made. National accounts are subject to revisions, and changing base years of PWT also introduces elements of non-comparability between different versions of PWT. Users have been advised of these changes with each update of PWT, but it is not always clear that the advice has been heeded. Figure 2 provides some evidence from a comparison of the latest two rounds of the PWT, versions 6.1 and 6.2. These show annual growth rates of real chained per capita GDP, computed over ten year periods, starting in 1960, 1970, 1980, and 1990, with 6.1 on the horizontal axis, and 6.2 on the vertical axis. Although most countries cluster close to the 45-degree line, there are a substantial number of revisions, particularly in the decade beginning in 1970, and particularly—but by no means exclusively—for African countries. As we might expect, revisions are greater for the poorer countries. The decade beginning 1950, and not shown here, had very little revision between the two rounds.

Ongoing work at IMF undertaken by Johnson, Larson, Papageorgiou and Subramanian (JLPS) sheds some light on the consequences of these revisions. At this stage of their work the authors would prefer that the specifics of their results not be quoted, so the following is a general summary of their findings thus far.

One issue is differences in growth rates between PWT versions. Annual growth rates of GDP based on the chain series were compared from PWT 6.1 and 6.2 for 40, 10 and annual intervals. Why would they differ? The discussion above suggests the main reasons: weights for growth of *C*, *I* and *G* change for both the fixed and chain indexes,

national accounts are revised and rebased, and new benchmark information becomes available. For all countries the annual cross section is fairly similar, the ten-year growth rates generally differ by less than 10 percent, but the 40 year rates often differ by 20 percent with some embarrassing outliers. Differences are lower for high-income countries, and larger for both low and middle-income countries.

PWT assigns quality grades to countries on an A to D standard based upon number of benchmark comparisons in which a country has participated and some internal measures of data stability over time. With the exception of Singapore, there is for practical purposes no difference in growth rates for the remaining 30 A and B grade countries. If users have employed PWT grades as variables or to group countries in their analysis, it has not come to our attention. In work with earlier benchmarks it was found that differences in growth rates in PWT and in national growth rates were largest for countries where their own national growth rates were affected by using a new base year.

Do these growth differences between versions of PWT make a difference? JLPS undertook an extensive literature search and performed a number of replications involving PWT 5.6, PWT 6.1 and PWT 6.2. The preliminary guidelines thus far are:

- SAFE: Studies that mainly use long-term growth, 40 year intervals, are fairly robust with respect to use of any of the above versions of PWT. This may also be true for 10 year intervals, or at least the conclusions of such studies do not appear dependent on which version of PWT is used.
- SAFE: Use of annual growth rates for the A and B grade countries is safe.
- NOT SAFE: It is not safe to use annual growth rates from different versions of PWT for non-OECD countries.

- NOT SAFE: Different versions of PWT are not robust for dynamic analysis at annual frequencies.

The IMF group plans to check whether PWT 7.0, which will incorporate the 2005 ICP, leads them to change their story. This brings us full circle to the problem of integrating the new view of the world economy in the 2005 ICP with the older view in the WDI and PWT. As in the past, the plan is to present a reasonable view of the world economy in 2005 and to move that backward and forward in time. The work of JLPS suggests that PWT should follow standard national statistical practice and provide a consistent set of accounts on a 2005 base in PWT 7.0, with full incorporation of the 2005 ICP. The JLPS research also suggests that more alternatives need to be considered for updating and backdating the PWT numbers through time. However, as noted, PWT is not designed for temporal analysis of annual data, and there are national data sets that can be employed by users that wish to do that type of analysis.

A recent paper by Ciccone and Jarocinski (2008) (CJ) identifies another type of analysis for which revisions in the PWT are a serious problem. These are attempts based on automatic model selection procedures to have the data choose which variables should appear in growth regressions. CJ apply these procedures to version 6.2 and 6.1 and obtain very different variables; in one of their experiments, the two data sets disagree on 13 out of 23 growth determinants, some of which are widely used in the literature. When they restrict the range of variables over which selection is possible, the results are a good deal more robust, if not perfectly so. These experiments may tell us as much or more about the failings of statistical model selection than about the failings of the PWT. Such procedures have are sensitive to small changes in the data that makes them unsuited for use with the

PWT, none of which implies that the data are not well-suited to answering better-structured or more theoretically informed questions.

For the reasons already discussed, the results of ICP 2005 will not be woven into PWT 7.0 without adjustment. For example, the fixity restrictions will not be respected, and some preliminary analysis suggests that this will make a marked difference to some of the PPPs, see again Table 2. Other possible adjustments are also under consideration, including modifications for the special character of Chinese prices, and for the lack of comparability of non-priced goods and services across the regions. What can be said is that because of ICP 2005 there is a much richer data set available for those researchers interested in differences of economic structure and income across countries than has been available until now.

5. Conclusions

This summary of PPP national accounts, and of the Penn World Table in particular, has covered only a fraction of the issues that go into the construction of these data. Experience suggests that it is hard to know in advance which features of the data are likely to be decisive for which purpose, or which particular detail will be responsible for some new or potentially interesting finding. Perhaps the overriding message it to exercise caution, particularly with comparisons between countries whose economies are very different, and particularly with the national accounts data provided by countries whose statistical capacity is weak. On the former, there are deep conceptual difficulties that cannot be resolved by collecting better data. On the latter, it must always be remembered that the international accounts are no better than the national accounts of the participating

countries. The quality ratings in the PWT contain useful information that should be more heavily used.

There are also some specific health warnings that are worth emphasizing. One is about index numbers, and the general point that price indexes are not prices. Although most economists know that different price indexes give different answers, the comfort that comes from thinking that it matters little in practice is a strictly domestic comfort that does not always travel internationally. Differences between Paasche and Laspyeres indexes are sometimes very large in the ICP, and these extend to differences between EKS and GK aggregations that are used by different agencies, see again Tables 1 and 2. Second, given the regional structure of the ICP, it is always worth taking into account the possibility that the regions are not comparable in some important respect. Regional dummies are often included in growth regressions for substantive reasons—or as an admission of ignorance—but there are also statistical reasons for including them, or even, when possible, treating the regions separately. Third, there are particular reasons for caution in using the data for countries of the CIS and the former Soviet Union. Fourth, some important components of GDP, including government services, health care, education, construction, and the rental of housing, are extremely difficult to compare across countries, and are often handled by sensible, but more or less arbitrary assumptions. Not only is it dangerous to rely on the benchmark estimates for these items, but their treatment can affect overall PPPs between countries, or even regions.

Researchers have a wide range of data sources. The World Bank's WDI contain time-series for GDP in current and constant international dollars, but not for the other components of the national accounts. Eurostat provides PPP accounts for Europe and the

OECD, with GK versions following the EKS data with a one year lag. Time series data go back to 1980 in some cases, and there are some data at higher than annual frequency. For researchers who do not need data from Asia, Africa, or Latin America, these are worth serious consideration. The PWT, on which we have focused, is essentially the only option for long time series containing the main variables of the national accounts.

We end with a horror story designed to serve as a last warning. Successive editions of the ICP have tended to revise upward the PPPs of poor countries relative to those of rich, even in the same year. Closer comparability of the goods and services to be priced is one of the reasons; most recently, more appropriate assumptions about the productivity of government workers has had the same effect. In particular, successive revisions have tended to make India and China poorer than the previous revision, though by no means at the same rate. For several years, in spite of China's much more rapid growth, the ratio of Chinese to Indian per capita GDP did not rise by as much as would seem warranted, leading to suspicions that the government of China wished simultaneously to exaggerate its growth rate and to understate its level of per capita GDP, see Srinivasan (1994, p.10). The recent (apparent) shrinkage of both India and China in the 2005 ICP has provoked similar concerns, most notably from Bhalla (2008), a long-time if not always reliable critic of the World Bank. Using the numbers here, a version of Bhalla's argument is as follows. The 2005 estimate of Chinese per capita GDP at 2005 international dollars is $4,091. According to the PWT6.2, which is based on the Chinese official data, China grew at 5.52 percent a year from 1952 to 2004; at this rate, GDP per capita in 1952 would have been $279 in 2005 international dollars, or $153 at 1985 international dollars converted using the US CPI. Pritchett (1997) has persuasively argued that approximately

$250 in 1985 international dollars is the minimum level of per capita GDP that is required to sustain a population, or that has ever been observed for more than a short period. If so, it is simply not possible that *both* the current PPP estimate of Chinese GDP *and* the official growth rates of the economy can be correct. On the latter, Maddison (2007) estimates Chinese GDP growth since 1952 at "only" 4.4 percent a year, but this still leads to $229 in 1985 international dollars, still below Pritchett's cutoff. Reducing the PPP by ten percent or so, as suggested in Section 3.2 above, would bring this number into an only somewhat more plausible range. More broadly, the point remains that many of these numbers have substantial uncertainty, and that extrapolations can easily lead to results that make no sense.

References

Afriat, Sydney, 1967, "The construction of utility functions from expenditure data," *International Economic Review*, 8(1), 67–77.

Afriat, Sydney, 1972, "The theory of international comparisons of real income and prices," in D. J. Daly, ed., *International comparisons of prices and output, Proceedings of the conference at York University, Toronto*, New York, National Bureau of Economic Research, 13–84.

Asian Development Bank, 2007, *Purchasing power parities and real expenditures*, Manila, December.

Aten, Bettina, 2006. 'Interarea price levels: an experimental methodology', *Monthly Labor Review*, 129(9), September, 47–61.

Bhalla, Surjit S, 2008, "World Bank—most Asians dead in 1950," *Business Standard*, August 23.

Brandt, Loren, and Carsten Holz, 2004, "Spatial Price Differences in China: Estimates and Implications," *Economic Development and Cultural Change*, 55(1), 43–86.

Chen, Shaohua, and Martin Ravallion, 2008, "China is poorer than we thought, but no less successful in the fight against poverty," Washington, D.C., Policy Research Working Paper 4621 (May).

Ciccone, Antonio and Marek Jarocinksi, 2008, "Determinants of Economic Growth: Will Data Tell?, Working Paper Series No. 852, European Central Bank, January.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers, 2006, "Missing in action: teacher and health worker absence in developing countries," *Journal of Economic Perspectives*, 20(1), 91–116.

Deaton, Angus, Vivi Alatas, and Jed Friedman, 2004, "Purchasing power parity exchange rates from household survey data: India and Indonesia," Princeton, Research Program in Development Studies.

Diewert, W. Erwin, 1976, "Exact and superlative index numbers," *Journal of Econometrics*, 4(2), 115–45.

Diewert, W. Erwin, 1999, "Axiomatic and economic approaches to international comparisons," in Alan Heston and Robert Lipsey, eds., *International and interarea comparisons of income, output, and prices*, University of Chicago Press, 13–52.

Diewert, W. Erwin, 2001, "The consumer price index and index number purpose," *Journal of Economic and Social Measurement*, 27, 167–248.

Diewert, W. Erwin, 2008, "New methodology for linking the regions," Dept. of Economics, University of British Columbia (September.) Available at http://www.econ.ubc.ca/diewert/dp0807.pdf

Dikhanov, Yuri, 2007, "Methodological note—PRC: estimating national average prices," Appendix 1 in Asian Development Bank, *Purchasing power parities and real expenditures*, Manila, ADB, pp 118–30.

Dowrick, Steve, and John Quiggin, 1994, "International comparisons of living standards and tastes: a revealed preference approach," *American Economic Review*, 332–41.

Eltetö, O., and P. Köves, 1964, "On a problem of index number computation relating to international comparison," *Statisztikai Szemle* 42, 507–18.

Feenstra, Robert, Alan Heston, Marcel P. Timmer and Haiyan Deng , 2009, "Estimating Real Production *and* Expenditures Across Nations: A Proposal for Improving the Penn World Tables", *Review of Economics and Statistics*, forthcoming.

Geary, Roy C., 1958, "A note on the comparison of exchange rates and purchasing power between countries," *Journal of the Royal Statistical Society, Series A (General)*, 121(1), 97–99.

Gershenkron, Alexander, 1947, "The Soviet indices of industrial production," *Review of Economic Studies*, 29(4), 217–26.

Gini, Corrado, 1924, "Quelques considerations au sujet de la construction des nombres indices des prix et des questions analogues," *Metron*, 4, 3–162.

Heston, Alan (2008) "The 2005 Global Report on PPPs: A Preliminary Review," *Economic and Political Weekly* (Mumbai), March 15, 2008, 65–9.

Johnson, Simon, William Larson, Chris Papageorgiou and Arvind Subramanian (2008) , "Some Health Warnings About Using PWT Growth Rates," Presentation. PWT Workshop, May 11th.

Khamis, Salem H., 1972, "A new system of index numbers for national and international purposes," *Journal of the Royal Statistical Society, Series A (General)*, 135(1), 96–121.

Kolkoski, Mary, Brent Moulton, and Kim Zieschang, 1999, "Interarea price comparisons for heterogeneous goods and several levels of commodity aggregation," in Alan Heston and Robert Lipsey, eds., *International and interarea comparisons of income, output and prices,* University of Chicago Press. 123–66.

Maddison, Angus, 2003, *The world economy: historical statistics*, Paris, OECD.

Maddison, Angus, 2007, *Chinese Economic Performance in the Long Run: Second Edition, Revised and Updated 960-2030 AD*, Paris, OECD. Appendix D.

Neary, J. Peter, 2004, "Rationalizing the Penn World Table: true multilateral indexes for international comparisons of real income," *American Economic Review*, 94(5), 1411–28.

Nuxoll, Daniel A., 1994, "Differences in relative prices and international differences in growth rates," *American Economic Review*, 84(5), 1423–36.

Pritchett, Lant, 1997, "Divergence, big time," *Journal of Economic Perspectives*, 11(3), 3–17.

Rogoff, Kenneth, 1996, "The purchasing power parity puzzle," *Journal of Economic Literature*, 34(2), 647–68.

Seers, Dudley, 1983, *The political economy of nationalism*, Oxford. Oxford University Press.

Srinivasan, T. N., 1994, "Data base for development analysis: an overview," *Journal of Development Economics*, 44, 3–27.

Summers, Robert, and Alan Heston, 1991, "The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988," *Quarterly Journal of Economics*, 106(2), 327–68.

Szulc, B., 1964, "Indices for multiregional comparisons," *Przeglad Statystyczny*, 3, 239–54.

Van Veelen, Matthijs, 2002, "An impossibility theorem concerning multilateral international comparisons of volumes," *Econometrica*, 70, 369–75.

World Bank, 2007, *World Development Indicators*, Washington, DC. (CD.)

World Bank, 2008a, *Global purchasing power parities and real expenditures: 2005 international comparison program*, Washington, DC.

World Bank, 2008b, "ICP 2003–2006 Handbook," available online at http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/ICPEXT/0,contentMDK:20126612~pagePK:60002244~piPK:62002388~theSitePK:270065,00.html

World Bank, 2008c, *World Development Indicators*, Washington, DC. (CD)

Table 1: Ratios of Laspeyres to Paasche price indexes, USA and Nigeria versus selected other countries

USA		NIGERIA	
Ten smallest			
Ireland	1.048	Congo	0.962
Canada	1.057	Latvia	1.003
Austria	1.062	Guinea	1.030
Switzerland	1.068	Mali	1.032
Germany	1.072	Zimbabwe	1.049
Luxemburg	1.080	Albania	1.050
Italy	1.091	Estonia	1.051
Austria	1.092	Equatorial Guinea	1.053
Belgium	1.094	Burkina Faso	1.055
UK	1.096	Chad	1.062
Ten largest			
Djibouti	2.798	Oman	1.674
Armenia	2.811	Korea	1.707
Moldova	2.905	Japan	1.747
Rwanda	2.943	USA	1.780
Burundi	3.055	Iran	1.781
Zimbabwe	3.771	Kuwait	1.945
Chad	4.033	Bahrein	1.978
Gambia	4.210	Luxemburg	2.095
Kyrgyzstan	5.107	Tajikistan	2.473
Tajikistan	9.615	Qatar	2.847
Other important			
China	1.658	China	1.375
India	1.611	India	1.390
Indonesia	1.484	Indonesia	1.405
Brazil	1.508	Brazil	1.329
Nigeria	1.780	South Africa	1.069
Russia	1.823	Russia	1.092

Table 2: Bilateral Fisher, EKS, and Geary-Khamis indexes, selected countries as in Table 1

	LP-Spread	Bilateral Fisher	EKS	Geary-Khamis	ICP 2005
Ten smallest					
Ireland	1.048	1.042	1.023	1.100	1.023
Canada	1.057	1.195	1.176	1.237	1.214
Austria	1.062	0.878	0.852	0.902	0.874
Switzerland	1.068	1.748	1.709	1.695	1.741
Germany	1.072	0.896	0.860	0.873	0.893
Luxemburg	1.080	0.938	0.924	0.854	0.922
Italy	1.091	0.884	0.853	0.890	0.875
Australia	1.092	1.375	1.319	1.401	1.388
Belgium	1.094	0.889	0.880	0.889	0.899
UK	1.096	0.637	0.649	0.639	0.649
Ten largest					
Djibouti	2.798	73.62	84.69	65.94	84.69
Armenia	2.811	145.15	147.07	123.53	178.58
Moldova	2.905	3.617	3.513	3.050	4.434
Rwanda	2.943	176.28	188.63	152.94	186.18
Burundi	3.055	315.61	343.34	276.49	342.96
Zimbabwe	3.771	26702	30671	21900	33068
Chad	4.033	163.04	205.92	143.0	208.0
Gambia	4.210	6.315	8.019	5.530	7.560
Kyrgyzstan	5.107	8.157	9.068	6.390	11.354
Tajikistan	9.615	0.457	0.542	0.296	0.744
Other important					
China	1.658	3.127	3.059	2.941	3.448
India	1.611	14.755	13.480	13.605	14.669
Indonesia	1.484	3833	3606	3540	3934
Brazil	1.508	1.346	1.376	1.401	1.357
Nigeria	1.780	54.77	57.42	54.38	60.23
Russia	1.823	11.314	11.163	11.397	12.736

Notes: US is the base country. LP-spread is the Laspeyres-Paasche spread as in Table 1. Bilateral Fisher is the Fisher price index for each country relative to the US, calculated as a bilateral comparison. EKS and Geary-Khamis are the multilateral PPP indexes for GDP excluding the trade balance. ICP2005 is the PPP from the WDI 2008.

Source: Authors' calculations using basic heading parities from the 2005 ICP.

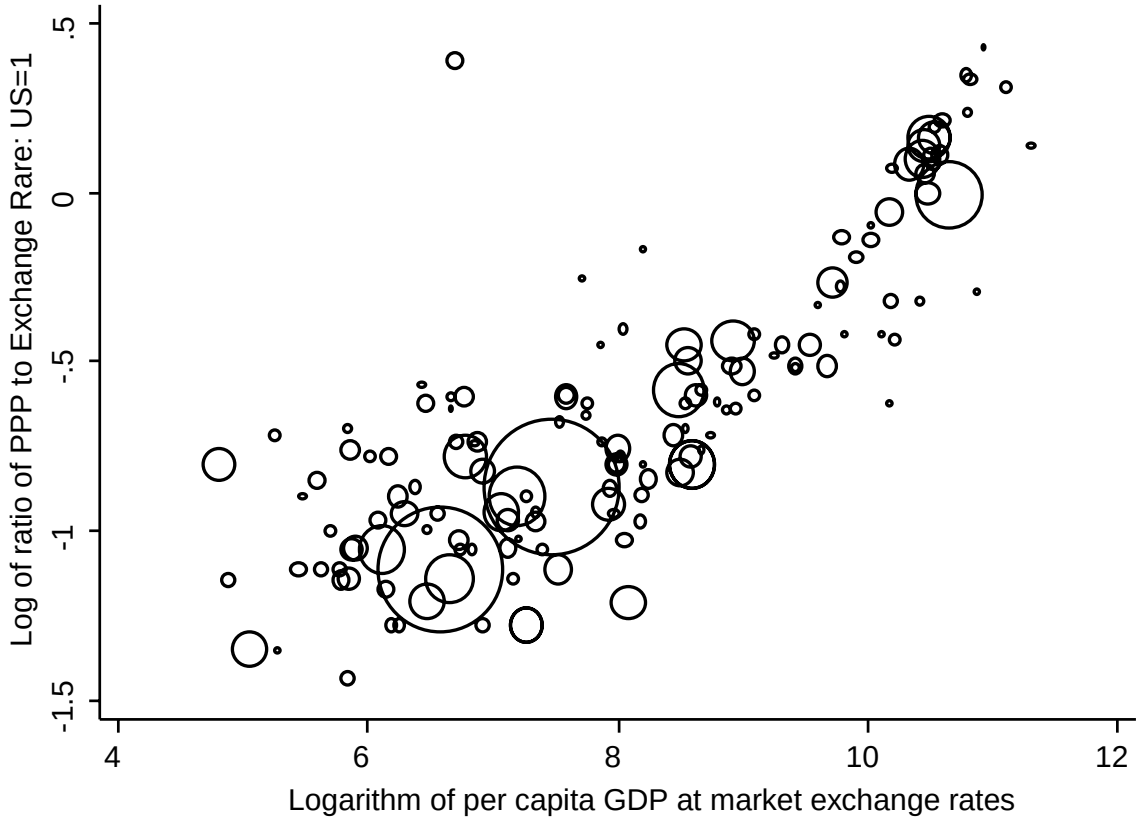

Figure 1: Price levels in relation to GDP

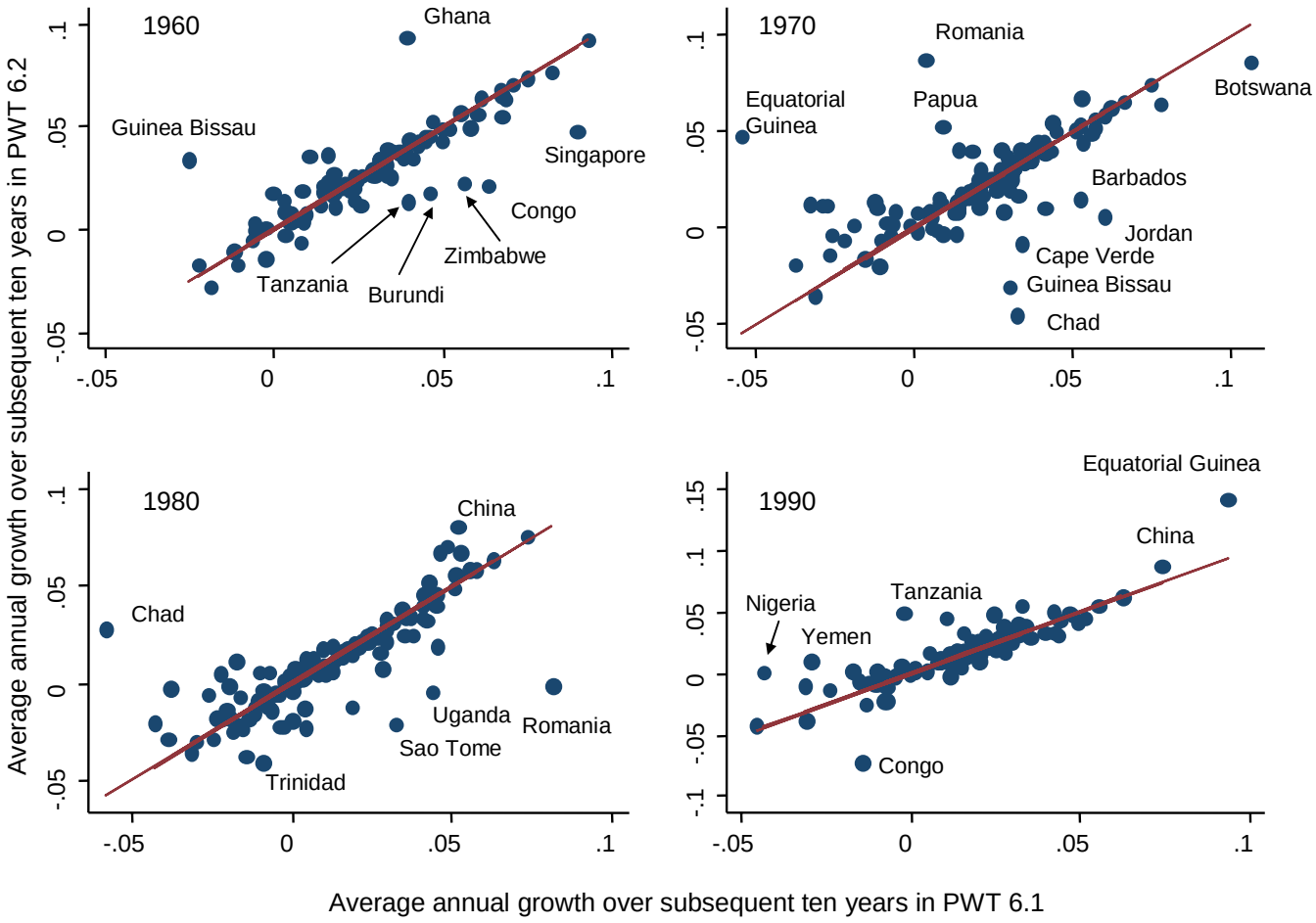

Figure 2: Revisions to ten year growth rates in PWT6.1 and PWT6.2