

NBER WORKING PAPER SERIES

TRYING TO UNDERSTAND THE PPPS IN ICP2011:
WHY ARE THE RESULTS SO DIFFERENT?

Angus Deaton
Bettina Aten

Working Paper 20244
<http://www.nber.org/papers/w20244>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2014

We are grateful to Orley Ashenfelter, Grant Cameron, Erwin Diewert, Rob Feenstra, Nada Hamadeh, Alan Heston, Robert Inklaar, Michel Mouyelo-Katoula, Paul McCarthy, Prasada Rao, Marko Rissanen, Marcel Timmer, and Fred Vogel for help, comments and suggestions. Both Aten and Deaton are consultants to the International Comparison Program (ICP); Aten as a member of the Computation Task Force, and Deaton as a member of the Technical Advisory Group. The research reported in this paper was not funded by the ICP. The data used in the paper are either from the World Development Indicators or can be accessed via application to the ICP; for details see

http://siteresources.worldbank.org/ICPINT/Resources/270056-1255977254560/121120_ICPDataAccessPrinciples&Procedures.pdf

The views expressed in this paper are solely those of the authors and not necessarily those of the U.S. Bureau of Economic Analysis, the U.S. Department of Commerce, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Angus Deaton and Bettina Aten. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Trying to Understand the PPPs in ICP2011: Why are the Results so Different?

Angus Deaton and Bettina Aten

NBER Working Paper No. 20244

June 2014

JEL No. E01,F00,O11,O47

ABSTRACT

The 2011 round of the International Comparison Program (ICP) has published a set of purchasing power parities (PPPs) that are sharply different from those that were expected from extrapolation of the 2005 round. In particular, the world in 2011 looks sharply more equal than previously calculated, because consumption and GDP in most poor countries were revised upward relative to the U.S. and other rich countries. Here we attempt to find out what happened. It is first noted that the 2005 round was itself sharply different from what was then expected, and made the world much less equal. We argue that the 2011 round is superior to the 2005 round, and that many of the changes in 2011 undo what happened in 2005. We identify a likely source of the problem, which is the way that the regions of the ICP were linked in 2005. We use two different methods for measuring the size of the effect. Both suggest that the 2005 PPPs for consumption for countries in Asia (excluding Japan), Western Asia, and Africa were overstated by between 20 to 30 percent. If these results are correct, they call for substantive backward revision of international comparisons, as well as estimates of global poverty and inequality.

Angus Deaton

361 Wallace Hall

Woodrow Wilson School

Princeton University

Princeton, NJ 08544-1013

and NBER

deaton@princeton.edu

Bettina Aten

Bureau of Economic Analysis

1441 L Street, N.W.

Washington, DC 20230

Bettina.Aten@bea.gov

1. Introduction

A summary of the results and findings of the 2011 round of the International Comparison Program was published on April 30th, 2014, World Bank (2014); the report gives purchasing power parity exchange rates for the main aggregates of GDP for 199 economies. Until the publication of the new results, the World Bank in its World Development Indicators provided extrapolated PPP exchange rates; these were based on the 2005 round of the ICP, updated using relative inflation rates for each country. The new estimates for 2011 from ICP 2011 are quite different from these extrapolations. In particular, most poor countries of the world are estimated to be larger relative to the US and other rich countries than was estimated from the extrapolations. This aspect of the results has attracted a good deal of attention in the press, particularly the fact that the aggregate Chinese economy is much closer to the US than previously estimated, and also that the Indian economy is now estimated to be larger than the Japanese economy. The new results also sharply reduce previous estimates of international inequality.

The PPPs for individual consumption by households, like the PPPs for GDP, have been revised downwards, so that estimated consumption levels outside the rich countries are now higher than previously estimated. As was immediately noted by several commentators, Dykstra, Kenney, and Sandefur (2014), and Chandy and Kharas (2014), these consumption PPPs are the relevant ones for the calculation of global poverty rates, which are likely to be deeply affected. If the global poverty line is held fixed in US dollars, for example at \$1.25 as currently used by the World Bank, its local value in poor country currencies will now be lower, so that there will be

fewer people living below it. These calculations show very large declines in estimated poverty rates for 2010—which is the last year covered by the World Bank’s poverty calculator. Chandy and Kharas calculate that the number of people living below \$1.25 in 2010 has fallen from 1,215 million using the PPPs extrapolated from ICP 2005, to 571 million using the new PPPs back cast from 2011 to 2010. Poverty in India falls by more than 300 million people.

The 2011 round of the ICP contained many methodological improvements over ICP 2005 but even in the absence of methodological changes, the results from a new round will not be identical to those extrapolated from an earlier round. Extrapolation is at best a short cut to the collection of new benchmark data. Even so, it remains unclear exactly why the 2011 results are quite so different from the extrapolations and, without an understanding, there will be continuing questions about whether the new results really are better than the old, and whether it is safe to use them in applications such as the measurement of global poverty and global inequality.

This is not the first time that a new round of the ICP has brought large changes to PPPs and to poverty counts. The 2005 round generated large *increases* in poor-country PPPs relative to the US compared with those that had been previously extrapolated from the 1993/95 round. These changes caused a sharp *upward* revision of the number of poor people below a global line denominated and held fixed in US dollars. There were also upward revisions in 1993/95 compared with the previous round in 1985; these also had large effects on the poverty calculations, Deaton (2001). When we think about what happened in 2011, we must keep this previous

history in mind. In particular, if 2011 is inconsistent with 2005, the problems could come from either ICP 2005 or ICP 2011, or both.

In this paper, we focus on the PPPs for individual household consumption that are used in the poverty estimates. This also simplifies the discussion by taking off the table some (although not all) of the most difficult (“comparison resistant”) items, such as construction or government services, as well as the trade-balance, which is treated differently in ICP from national accounts, and so predictably causes discrepancies between ICP estimates of GDP and extrapolations based on national accounts, McCarthy (2011).

In section 2, we start by documenting the changes in the PPPs in 2011 and look back on the corresponding changes in 2005. These calculations demonstrate that, at least to some extent, the 2011 ICP undid some of the changes that happened in the 2005 ICP. In section 3, we discuss three possible explanations for the conflict between extrapolations and actuals, two general, and one specific. The first general point is that extrapolation cannot be expected to match actual changes, essentially because the ICP is based on international multilateral price indexes that are different from the national price indexes that are used for extrapolation. The second general point is that PPPs have large, but largely unrecognized, standard errors, so that at least some of the changes with each round likely come from that source. Finally, and more specifically, we check whether there is evidence that there is a problem with the national prices indexes used for extrapolation.

In section 4, we examine the most likely explanation of the discrepancy, which is the change in the way that the regions were linked in 2011. The ICP is or-

ganized on a regional basis, with each region calculating a set of within-region PPPs. In order to get a world table, these regions have to be linked. In 2005, this was done using a “ring” of eighteen countries, distributed over the regions, for whom there was essentially a separate ICP exercise. The PPPs for those countries were then used to link the regions. At the time of the 2005 ICP, there were concerns about the implementation of the ring method—for example, some of the within-region ring comparisons were not consistent with the main ICP, and some items on the ring list were difficult to price in some of the ring countries—and the ring method was not used in 2011. Instead, a global core list was developed, and *all* countries priced (subsets of) the items on the list. This allowed all calculations to come from a single set of data, and *all* countries in all regions contributed to the regional linking in a symmetric way.

Deaton (2010) suggested that the ring in ICP 2005 may have contributed to over-pricing in poor countries relative to rich countries, so that it is possible that ICP 2011, with its global core list, undid this overpricing. We provide some evidence from the ring countries that is consistent with this hypothesis. If this is correct, the main problem lies with 2005, not 2011. Our estimates suggest that countries in the African, Asian, and Western Asian region of the ICP had consumption PPPs that were overestimated by 20 to 25 percent in 2005.

Section 4 also provides an alternative approach based on an examination of the relationship between price levels and per capita consumption. According to the Balassa-Samuelson theorem, price levels—the ratios of PPPs to exchange rates—are higher in richer countries and rise as countries become richer. Ravallion (2012)

suggested that the relationship between price levels and per capita consumption is stable enough to use for extrapolation and gives a better fit than CPI extrapolation, a finding that was challenged by Inklaar (2013). Here, we use the relationship for a different purpose, to assess whether price levels in 2005 were systematically overstated. Our estimates from this method are consistent with those from examination of the ring, and once again suggest a 20 to 30 percent overestimation in Africa, Asia, and Western Asia in 2005.

Our results provide a plausible story for the conflict between the 2005 and 2011 ICP benchmarks (as well as for the conflict between the 1993/96 and 2005 benchmarks), but the evidence is far from conclusive. There are many other changes from 2005 to 2011 that the ring hypothesis does not explain; given the many other things that changed, that is not in itself surprising.

2. What happened?

It is impossible to understand what happened in the 2011 ICP without going back to the comparable results from the 2005 ICP, and how those in turn differed from the extrapolations based on the previous (1993/95) round. Figure 1 plots the ratio of the “new” (i.e. ICP 2005 PPP) to “old” (i.e. extrapolated from 1993/95) PPPs for individual household consumption in 2005 against the logarithm of per capita GDP in 2005 international prices. All data relate to 2005 and the GDP figures are those taken from the report of the 2005 round of the ICP. The immediate feature of this figure is that the 2005 revisions were mostly upward, i.e. the ratios are greater than one. The countries on the right that have a ratio of close to one are the Eurostat coun-

tries, who have an annual program, and therefore only minimal revisions at new rounds of the ICP. Otherwise, there is a pronounced negative slope, so that the lower was per capita GDP, the larger the upward revision in the consumption PPP. Relative to the US, poor countries got poorer or, equivalently, relative to most of the world, the US and other rich countries got richer; either way, the 2005 ICPs made the world distribution of income much more unequal compared with extrapolations from the previous round. Some of the revisions are very large, mostly in sub-Saharan Africa, but there were also upward revisions in Asia, including factors of 2.6 for Philippines, 2.3 for China, 2.1 for Indonesia, 1.9 for Nepal, and 1.4 for India and for Bangladesh.

When the 2005 results were published, these revisions were noted, but at the time, there seemed no pressing need to explain them. ICP 2005 had introduced many methodological improvements compared with previous rounds, and there had long been a widespread concern about the quality of the 1993/95 ICP. Many new countries had been surveyed for the first time in 2005, including many in sub-Saharan Africa where the largest revisions had taken place. China officially participated in the ICP for the first time in 2005 and India did not participate between 1978 and 2005. It was therefore reasonable to conclude that the new results were simply better, and that the revisions were largely a consequence of the poor quality of the 1993/95 round.

Figure 2 replicates Figure 1, but compares ICP 2011 with extrapolations based on ICP 2005. It shows the ratio of consumption PPPs from ICP 2011 to the extrapolations for 2011, plotted against per capita GDP in 2011 international prices. Figure 2 is in some respects the opposite of Figure 1. The ratios of new to old are

now mostly less than one instead of greater than one. Just as countries got poorer relative to previous calculations in the 2005 round, they got richer in the 2011 round. Even so, the pronounced negative slope in Figure 1 does not turn into a positive slope in Figure 2. But the graph is somewhat misleading because it includes, on the bottom right, a number of wealthy non-OECD countries, most of which are oil producers and whose per capita GDP levels are thereby inflated relative to consumption. If those countries are excluded, the positive slope is apparent. As we shall see in section 3 below, perhaps the most important methodological improvement in 2011 over 2005 is capable of explaining these results, not country by country, but for regions as a whole.

Figure 3 matches up the revisions in 2011 with the revisions in 2005 for the countries that were covered in both rounds. The regression line is constrained to go through the zero point, has a slope of -0.4 with a t -value of -9.9 , and shows that, on average, the countries that had the largest upward revisions to their consumption PPPs in 2005 relative to extrapolation had the largest downward revisions relative to extrapolation in 2011. There is much scatter around the line, so it is *not* true that ICP 2011 simply undid the revisions in ICP 2005 on a country-by-country basis. Because the slope is -0.4 , not -1 , the consumption PPPs in 2011 are higher on average than if the 1993/95 round had been used for direct extrapolation to 2011 with the 2005 ICP ignored; only part of the increase in poor countries' consumption PPPs that happened in 2005 was undone in the 2011 round. That there is a large scatter in Figure 3 should not be a surprise, if only because of the known inadequacies of the 1993/95 round, because of the major methodological improvements introduced

in 2005, and the smaller, but still important, improvements in 2011. All of these changes are wrapped into this figure.

Table 1 shows the revisions for each of the ICP regions, again covering only those countries that were covered in both rounds. The revisions in 2005 relative to extrapolations were very large, particularly in Africa, though all of the rest of the world had higher PPPs relative to the rich countries. These changes made the world distribution of consumption appear to be radically more unequal. Depending on the region, about 40 percent of those increases were undone in 2011, though in Western Asia, the downward revision in 2011 was larger than the upward revision in 2005. The reduction in Africa in 2011 undid much less of the 2005 increase than was the case in India, so that, just as the ICP 2005 “Asianized” poverty, the ICP 2011 will “Africanize” it, at least if no offsetting changes are made to the global poverty line; this effect is amplified by the fact that there are many more people close to the global poverty line in South Asia than is the case in Africa.

Table 1 also shows the standard deviations of the log ratios of extrapolations to actuals. Apart from the CIS—among which there are many special problems—there is a marked reduction in dispersion in 2011 compared with 2005. This is what we would expect from the fact that the methodology and country coverage was much more stable between 2005 and 2011 than before 2005. A much larger share of the changes between 2005 and 2011 comes from between regional revisions than from within regional revisions. This does not tell us which of 2005 and 2011 is correct, only that, at least within the regions, they have more in common than with previous rounds. The finding is also consistent with our hypothesis that the main rea-

son for the differences between 2005 and 2011 is to do with the way that the regions were linked.

3. Why are new ICPs not consistent with extrapolations?

There is a long literature on why spatial and temporal price indexes cannot be consistent with one another, so the fact that revisions will occur at each ICP round should not be a surprise, see McCarthy (2012) for a recent summary and cites to the literature. Indeed, if extrapolations were consistently accurate, there would be no need for the ICP at all. Extrapolations, although intuitively attractive, cannot be expected to reproduce the numbers that come out of an ICP benchmarking exercise.

The PPPs from the International Comparison Program are *multilateral* price indexes, in which the PPP for each country depends on price relatives and weights from all of the countries in the system. This would be true even if there were only two countries in the world, and this simplest of cases can be used to illustrate the issues involved.

When the World Bank estimates a PPP by extrapolation, it starts from a baseline consumption PPP, for example from ICP 2005, and then updates using consumer price indexes. If the PPP for India in USD were 20, say, and if CPI inflation in both countries were the same, the extrapolated PPP would remain at 20. If Indian CPI inflation were higher than CPI inflation in the US, the extrapolated figure would be higher than 20, and if the Indian inflation were lower than inflation in the US, it would be lower than 20. Of course, domestic CPIs use only local prices and local weights, whereas the PPP, when it is calculated by the ICP, will combine weights for

both India and the US in calculating a PPP between them. Multilateral price indexes start from bilateral indexes between pairs of countries, and if those bilateral indexes are to satisfy minimal properties, for example, that the Indian to US PPP must be the reciprocal of the US to Indian PPP, then they must use weights from both countries. As a result, the change in the PPPs by extrapolation will be different from the changes from one ICP to the next.

In the two-country case, with countries 1 and 2, one simple formula for the relative change in CPI in 2 relative to 1 is

$$d \ln P_2 - d \ln P_1 = s_2' d \ln p_2 - s_1' d \ln p_1 \quad (1)$$

where P_2 and P_1 are the two country price indexes, p_2 and p_1 are the underlying vectors of prices of goods and services, and s_2 and s_1 are the two corresponding vectors of weights or budget shares. The PPP for 2 in terms of 1, by contrast, uses weights for both countries, so if we use a Törnqvist index, we have

$$\ln PPP_2 = 0.5(s_2 + s_1)'(\ln p_2 - \ln p_1) \quad (2)$$

If we combine these two to see what happens over time, we have, ignoring changes in weights,

$$d \ln PPP_2 = (d \ln P_2 - d \ln P_1) - 0.5(s_2 - s_1)'(d \ln p_2 + d \ln p_1) \quad (3)$$

The last term is the discrepancy between the extrapolation, which is the first term on the right, and the change in the PPP on the left hand side. Its magnitude will depend on the changes in the underlying relative prices, as well as differences in the consumption patterns in the two countries. The discrepancy is zero if there are no relative price changes in either country, if the structures of spending are the same,

or if the differences in spending structures are uncorrelated with average changes in relative prices. For countries with very different patterns of consumption, or with different structures of domestic inflation, the discrepancy could be large. For example, if the relative price of food rises around the world—as was the case from 2005 to 2011—the change in the benchmark PPPs for a poor country relative to a rich country could be substantially less than the extrapolated value. This line of enquiry is worth pursuing further, but preliminary calculations suggest that the effect was not large relative to the extrapolation discrepancies discussed here.

These results assume that the prices measured in the CPI are identical to the prices measured in the ICP, which is far from the case in practice. Cross-country indexes must match goods that are (a) identical in both locations, (b) reasonably commonly consumed in both places, so that the comparison lists for the ICP are usually quite different from the comparison lists of the CPI. Once again, it is hard to assess what this would contribute to the discrepancies.

Another concern is the quality of domestic CPI indexes, which varies considerably across countries; common problems are outdated weights, or a coverage that may be restricted to urban consumers, or even only to those who live in the capital city. CPIs are also relevant when thinking about why price level indexes (PLIs) differ between poor and rich countries, and how PLIs can be expected to change over time. CPIs cover *all* prices, including both traded and untraded goods, as opposed to exchange rates which are determined by flows of traded goods; indeed, the fact that exchange rates do not reflect the prices of untraded goods is the principal reason for collecting PPP data in the first place. As countries get richer, and wages rise, the

price of untraded goods, which are initially much lower than in poor than in rich countries, will rise, and the PLI will rise too, as all prices converge to world prices, see Obstfeld and Rogoff (1996, pp. 210–4) for a textbook account. Because the CPI covers both traded and untraded goods, and provided it is correctly measured, it will automatically rise relative to the exchange rate as the country grows richer. Ravallion (2012) calls this the dynamic Penn effect in a reference to the original Penn effect, by which poorer countries have lower prices relative to their exchange rates. If the CPI is correctly measured, extrapolation forward from a previous PPP should automatically capture any such effect, at least up to the other reservations already discussed in this section. The dynamic Penn effect does not invalidate the forward extrapolation of PPPs using CPIs.

Between 2005 and 2011, there was a large worldwide increase in the relative price of food in world markets; the FAO food price index doubled in nominal terms from 2005 to 2011, and rose by 50 percent in relative terms. In countries whose CPI weights are outdated, the share of food in the index will be too high—provided the country is growing—and the measured CPI growth will be too high. Because the extrapolation is done using CPIs, and if, as is plausible, CPI weights are more likely to be current in rich countries, the extrapolation would lead to consumption PPPs that are too high for poor countries relative to rich countries, which would be revealed at the time of a new PPP.

Without an enormous amount of detective work, it is hard to assess this possibility directly, but some insight can be given by comparing inflation rates from CPIs with inflation rates from the implicit price deflators of consumers' expenditure

in the national accounts. This is not perfect, because the coverage of the CPI and consumption in the national accounts will not be the same, but the implicit price deflator—current price divided by constant price consumption—is a current weighted Paasche index, while CPIs are generally Laspeyres indexes, so the effect of rising relative prices of food overstating the latter is likely to be indicated by a divergence in the two rates of inflation.

There are some countries where, indeed, CPI inflation between 2005 and 2011 was considerably higher than inflation in the price deflator. India is one example, where CPI inflation was 50.4 percent and deflator inflation was only 38.3 percent. But this is not generally the case. Chinese CPI inflation was 19.6 percent, nearly 11 percent *lower* than the 30.4 percent inflation in the deflator. The World Development Indicators contain 115 countries with both indexes and, of these, only 61 have CPI inflation greater than deflator inflation. There are two countries, Nicaragua and Mauritania, where CPI inflation exceeds deflator inflation by more than 0.50, and three countries, Venezuela, Tajikistan, and Tanzania where the deflator exceeds CPI inflation by more than 0.50. The remaining countries are shown in Figure 4, which plots the difference against the logarithm of per capita GDP in (the base year) 2005 in 2005 international dollars.

The Figure does not suggest that the food price story is the reason that extrapolation gave PPPs that were consistently too high, if only because the points are not consistently above the zero line for the poorer countries. Yet the wide divergence between the two rates of inflation, outside of the rich countries, gives serious concern about the quality of price data in many countries. And even if these prob-

lems cannot explain the fact that PPPs were consistently lower than extrapolated, they surely contribute to the variance that is observed in Figure 2 and Table 1.

A final issue that is not always given sufficient recognition is that PPPs are themselves subject to substantial uncertainty in a world where relative prices and consumption patterns differ substantially across countries. Deaton (2012) shows the ratio of the logarithm of the Laspeyres to the Paasche index is a measure of uncertainty about PPPs, and derives formulas for uncertainty in multilateral indexes. These calculations suggest very large margins of uncertainty, for example of 20 to 30 percent for US to India or China comparisons. These are derived for a single international exercise, and the errors in changes may be smaller if the sources of error are correlated over time. Yet it should be borne in mind that we can expect substantial round-to-round changes in the estimation of PPPs, even in the absence of real changes in the underlying economies.

4. Regional linking, the ring and the comparison between 2005 and 2011

4.1 Regional linking methods in 2005 and 2011

One of the most important methodological differences between the 2005 and 2011 rounds of the ICP was the way in which the regions were linked to give a full global set of international PPPs. The ICP is organized into regions, each of which runs its own set of price comparisons to calculate PPPs for each of the countries within its region. Each of these regional ICPs has its own numeraire so, at this first stage, it is not possible to compare, for example, the United States, which is in the Eurostat/OECD region, with Cameroon, which is in the African region. At a final stage, a

set of price indexes is calculated, one for each region, and these are used to scale the regional parities, respecting the original relative PPPs within regions, and to reach a full international comparison. It is this last step that was done differently in 2011 from 2005. Because linking moves entire regions, it is an obvious place to look if we suspect that all of Asia, or all of Africa, was too low or too high in one of the rounds.

In 2005, regional linking was done using eighteen “ring” countries, selected so that there were at least two in each region; these are listed in the first column of Table 2. The ring methodology was never seen as ideal, but was a response to a number of constraints in the 2005 round, some of which were not apparent at the planning stage. By the time of the linking stage, most of the regions had completed their own intraregional PPPs, and those had to be respected in the global analysis. The selection of ring countries was also based on practical considerations, given that not all countries were prepared to undertake a second, different, pricing exercise.

The ring worked through a separate price comparison exercise for the ring countries alone, ignoring the rest of the world, and using a specially constructed list of more than 1,000 goods and services. These PPPs for the eighteen ring countries were then used to derive regional price indexes that could be used for linking. The details of the method, and the formulas, are given in Diewert (2013) or in Deaton and Heston (2010). It starts by selecting a numeraire country in each region, and then converting all the regional prices into the currency of that region’s numeraire. In this way, each region becomes a “super-country” to be linked using the regional price indexes from the ring. After the round was complete, it was discovered that the super-country method is not invariant to the choice of numeraire country within

each region; it is unlikely that this caused much error compared with possible alternatives, but it certainly contributed to the decision to abandon the ring in 2011.

When the 2005 ICP was completed and reviewed, a number of other problems with the ring method became apparent. Some of the ring countries were very different from one another, for example Cameroon and Japan, or Brazil and Oman, making comparisons difficult, and raising the concern that the selection of ring countries might be having a large effect on the PPPs of all countries. Britain is the only representative of Western Europe (or, apart from Japan, of the OECD), while there are four representatives in sub-Saharan Africa. Beyond that, some of the within-region comparisons from the ring were inconsistent with the comparisons from the previously constructed within-regional comparisons. The latter, which used regional lists adapted to each region's pattern of consumption, and which covered all countries in the region, were almost certainly superior to the ring comparisons, which were based on a common list, with some items hard to price, and for only a handful of countries in the region. Yet it was only the ring comparisons that determined the "tectonic" price indexes that linked one region with another.

Given all this, ICP 2011 used a common global "core" list, developed in consultation with all of the regions. Countries in each region priced as many of the items on the list as they could together with additional items selected by the region as suitable for their countries. Although not all countries would be able to price all the items on the core list, there was enough overlap so that, at the final linking stage, all countries would be involved, there would be many links across regions, and the goods and prices would be the same as those that had been used in the construction

of the within-region PPPs at the first stage. The precise linking method in 2011 is the Country Aggregation with Redistribution (CAR) method, first introduced by Kravis, Summers, and Heston (1982) in the 1980 ICP; a brief description is given in the Summary Report, World Bank (2014). The advantage of the CAR method over the ring lies in its consistency with the regional pricing strategy and in its incorporation of *all* countries into the linking of the regions so that we escape the dependence of the all-important regional linking factors on the choice of ring countries, or on the selection of goods and services for the ring list.

4.2 Did the 2005 ring overstate price levels in poor countries?

That the ring method might have overstated PPPs in poor countries was discussed in Deaton (2010) though, in the absence of the 2011 ICP, there were no decisive results. The ring list was extremely detailed, with a large number of items that were only widely available in rich countries. Enumerators in Cameroon, Kenya, Senegal, Zambia and Sri Lanka had to price a 2003 or 2004 vintage bottle of Bordeaux, front-loading washing machines with a pre-specified spin speed, and a Peugeot 407 with air conditioning and climate control. The obvious danger is that, when such items can be found at all, they are rare and expensive, or perhaps only available in stores patronized by a small subset of the population; in the absence of expenditure weights for items within basic heads, there will be an overstatement of the price level in poor relative to rich countries.

A second and related problem with ring pricing is what happens when a whole category of consumption is very expensive but very rarely consumed in the

poor country, but relatively cheap and widely consumed in the rich country. Air travel is an example that is relatively very expensive in many African countries. It has a low budget share in the African ring countries, so that its high price there has little effect on the overall domestic price level. But when we compare air travel in Cameroon, say, with air travel in Britain, say, the high relative price in Cameroon is weighted, not by the low Cameroon share, but by an average of the shares in Britain and Cameroon. This makes Cameroon look very expensive, even though hardly anyone there buys the good, and given the small number of countries in the ring, this high price of Cameroon relative to Britain can have a substantial effect on the price index for Africa relative to the OECD.

Questions of item specification—which can be thought of as quality issues—and of how to weight—are not specific to the ring method, but are inherent in making multilateral price comparisons across widely different countries. Yet it is plausible that the 2005 ring, with its very detailed list with many rich-country items, and the limited number of countries, led to an overestimation of price levels in the poorer countries in the ring, and thus, overestimated price levels for whole regions, particularly for Asia and for Africa.

We investigate the “ring overstatement hypothesis” by looking at how the ring countries fared in both 2005 and 2011. We calculate price levels for consumption for the ring countries using ICP data for 2005 and 2011, ignoring all other countries, so that we have comparable ring ICPs for both rounds; the 2005 version is the same as that used to link the regions, while the 2011 version is only for the purposes of this paper. We then use the two comparisons to calculate an ICP-based infla-

tion rate from 2005 to 2011 for each ring country, which we can compare with the inflation rate from the CPI. If the ring led to price overstatement in poor countries in 2005, the ICP-based inflation rates will be systematically lower than CPI inflation in poor countries, at least up to the noise that is apparent in Figure 4.

For 2005, we used the ring prices and ring list to calculate a set of PPPs, as if the ring countries were the only countries in the world. In 2011, we have done the same thing with the global core list, but again only using information from the eighteen countries that comprised the ring in 2005. Nothing from any other country has any effect on these numbers. For both rounds, we used the country-product dummy (CPD) method below the basic head level with weights reflecting surveyor-reported importance of each item (for 2011, unweighted in 2005), and expenditure-weighted CPD above it. Ideally, we should have liked to calculate price indexes in local currency for each country and each basic head, so as to make a direct comparison between 2011 and 2005. However, for many basic heads there is little overlap between the ring list in 2005 and the core list in 2011, so this approach is not feasible.

Table 2 lists the results. The PPPs and the exchange rates are scaled so that the UK is the numeraire. The relative inflation rates are changes in the logs of the CPIs from 2005 to 2011 from the World Development Indicators minus the corresponding inflation rate of the UK. We have adjusted the Estonian and Slovenian exchange rates and PPPs for 2011 back to their original currencies by multiplying by the rate at the time they joined the Euro.

If extrapolation were to work perfectly, the change in the logs of the PPPs from 2005 to 2011 would be equal to the inflation rate in the last column. We know

that there are many reasons why this is not going to happen. Even so, if for some countries the ring prices were overstated in 2005, but not in 2011, we would expect the change in the PPPs to be systematically *less* than the inflation rate because, in addition to true inflation, the overstatement would be unwinding. We think of the log change in the PPP as the predicted or ICP-based inflation rate, to be compared with the actual inflation rate, and what we are looking for is that the ICP-based inflation rate will be too low in the poorer countries of the ring, and approximately correct in the richer countries.

Figure 5 plots the ICP-based inflation prediction, the log of the 2011 PPP minus the log of the 2005 PPP, against the actual inflation rate. The graph provides a close to clean separation of the ring countries into two groups, a “rich” group, consisting of Japan, Chile, Hong Kong, Britain, Brazil, Estonia, Slovenia, and South Africa, and a “poor” group, consisting of Malaysia, Senegal, Cameroon, Philippines, Oman, Jordan, Zambia, Sri Lanka, Egypt, and Kenya. The figure shows the regression lines for the two groups; ideally these would have a slope of one, but they are substantially lower, 0.56 for the top group, and 0.58 for the bottom group, both of which are significantly different from unity. They can perhaps be explained by the substantial measurement error in inflation rates that we have already seen. If we pool the countries, and run a single regression with an intercept difference, the estimated difference is 0.26, which is the gap between the two lines in Figure 5. The standard error of the 0.26 estimate is 0.044, though this is misleadingly small, given that the choice of countries is itself based on inspection of the graph.

While we should resist the temptation to read too much into Figure 5, particularly given that there are only eighteen points, and should also remember the danger of overfitting, the figure is certainly consistent with the idea that the PLIs in the ring were too high in the poor group, but not in the rich group, as would be expected from the difficulty of pricing ring list items in the poorer countries. The exceptions are also interesting, with Japan below the top line, which would make sense if the ring list items were hard to price there, and South Africa well above the bottom line, which is what we would expect if the ring items were commonly purchased by the relatively wealthy white population. Cameroon and Zambia are the two countries with the lowest ICP inflation relative to CPI inflation, and they are countries where the ring items were likely the most unrepresentative and the hardest to find.

Another way of looking at the numbers in Figure 5 is to subtract the inflation rate calculated from the CPI from the inflation rate from the ICP. This is the discrepancy from the extrapolation. Given Figure 5, this constant can be roughly read off by looking at the rich countries; it is, of course zero for GBR, and runs from 0.07 for Chile to -0.23 for Estonia. For the poor countries, where the CPI increase is much larger than predicted from the ICP, the differences are more negative than this, running from -0.31 for Senegal to -0.64 for Zambia. If it was harder to find the ring list items the poorer was the country, we should expect to see a relationship between these residuals and the level of per capita GDP. Figure 6 shows that this is indeed the case.

We note again that there is a lot of variation around the line, but, subject to country to country variation in ring pricing in 2005, country to country variation in

the accuracy of the CPI, and all the other reasons why PPPs are not expected to inflate at the same rate as the CPI, it is perhaps surprising that the relationship in Figure 6 is as strong as it is. But it surely provides support for the idea that prices in the poorer countries in the ring were overstated in 2005.

If the analysis here is correct, and given that the ring prices moved whole regions relative to OECD/Eurostat, the \$ PPPs for Asia, Africa, and Western Asia were all too high in 2005. However, the 26 percent difference between the two lines is too large an estimate of the discrepancy, if only because, in Figure 5, South Africa (ZAF) is included with the rich countries, but was one of the African ring countries in 2005. If the lines are re-estimated with South Africa in the poor group, the estimated difference falls to 0.23, with a standard error of 0.0625. This number should be compared with the revisions in the second column of Table 1; if the 2005 ICPs for the three regions were reduced by this amount, the extrapolations to 2011 would have been close, not on a country-by-country basis, but region by region.

4.3 An alternative crosscheck calculation

A different and independent approach is to look for evidence of overstatement in the cross-country relationship between the price level index (PLI)—the ratio of the PPP to the exchange rate—and per capita consumption. The Balassa–Samuelson theorem posits that PLIs are lower in poor countries, and if this relationship is stable over time—which may or may not be the case—and if the PPPs of Africa, Asia and Western Asia were overstated in 2005, we should see an upward shift in the relationship for those regions in 2005.

Suppose that the relationship between the PLIs and per capita consumption is approximated by an equation that is linear in logs

$$\ln PLI^* = \alpha + \beta \ln C^* + u \quad (4)$$

and that this holds both across countries and across time, at least for the correctly measured C^* (in PPP terms) and PLI^* . Suppose that for three regions, Western Asia, Africa, and Asia, and in one year only, 2005, the PPPs are overstated, so that

$$\ln PPP = \ln PPP^* + \theta \quad (5)$$

where θ is a positive number for the selected regions in 2005, and zero otherwise.

If we substitute (5) into (4), we get the relationship between the measured PLI and measured C , which is

$$\ln PLI = \alpha + \theta(1 + \beta) + \beta \ln C + u \quad (6)$$

so that equation (6) should show an intercept shift in 2005 for the three regions.

Figure 7 starts by showing the scatter of the logs of the $PLIs$ and log per capita consumption in 2011, where the red circles are the 2011 ICP and the green circles are the extrapolations for 2011 based on ICP 2005 and the CPIs; the areas of the circles are proportional to population. The fitted lines to both actuals and extrapolations slope upward, and the latter is above the former. The figure also shows that the best fit is not in fact a straight line, and that the relationship is flatter among the poorer countries. This is convenient for our purposes, because we can fit a straight line in logs to the three affected regions, impose the same slope, and estimate the difference in intercept which turns out to be -0.325 which, according to equation (6), must be divided by 1.10 (one plus the slope) to give -0.29 , which is consistent with the second column of Table 1. Alternatively, we drop the assumption of an

identical slope, fit two straight lines, one for the extrapolated numbers and one for the actuals, and then calculate the average distance between the two lines for each country in the three regions; this calculation gives -0.327 , confirming the harmlessness of the equal slope assumption.

Figure 8 shows relationships between the logarithm of the *PLIs* and per capita consumption in ICP2005 and ICP2011. Both ICPs show the upward slope, but it is clear that, on the left of the figure, among the poorer countries, the green circles tend to be above the red circles at the same level of per capita consumption. Note too that there is a good deal of scatter; for example, although China got richer and its *PLI* increased, India, which also got richer, saw its *PLI* decline. Of course, if the ring overstatement hypothesis is correct, both of those changes should be larger because their *PLIs* in 2005 were too high. Once again, we can estimate a linear regression for countries in Western Asia, Africa, and Asia, imposing the same slope, to estimate a difference in intercept of -0.230 . Given that the slope of the regression is 0.146 , this gives an estimate of the *PPP* overstatement in ICP 2005 for these regions of -0.201 , see again equation (6). Once again, we have repeated the calculation without assuming equal slopes to get an -0.225 , almost identical to -0.230 .

From the previous subsection, we have a number of -0.23 for ring overstatement; here we have -0.29 for the extrapolation versus actuals in 2011, while a direct comparison of 2005 and 2011 leads to an estimate of -0.20 . If CPI extrapolation itself contributes to the discrepancy—either through CPI overstatement, or through the compositional effects in equation (3)—it would perhaps explain why the original -0.20 gets magnified to -0.29 . None of these calculations is very precise,

and all make assumptions about functional forms and the absence of other effects. Even so, all of our evidence is consistent with overestimation in 2005 for Asia, Africa, and Western Asia, by an amount between 20 and 30 percent.

5. Summary and conclusions

We began by documenting the differences between, on the one hand, the PPPs for consumption that were measured in the 2011 round of the ICP and, on the other hand, their extrapolations based on the 2005 round together with relative rates of inflation using country consumer price indexes. While ICP 2005 revised upward the PPPs for poor countries relative to the rich countries, ICP 2011 revised them downward. Country-by-country, the revisions in 2011 are only weakly correlated with the comparable revisions in 2005, and on aggregate, only remove part of them. In 2005, when there were large methodological differences for many countries, the discrepancies with the extrapolations varied both across regions and within them. In 2011, by contrast, where the methodological revisions were more limited, the biggest revisions were across regions, not within them.

The pattern of raising poor region PPPs in 2005, and lowering them in 2011 raises the possibility that there was a one time problem in 2005 that was, at least to some extent, undone in 2011. One possible explanation for this is that the regions were linked in 2005 using a group of eighteen ring countries, a method that was replaced by a more robust and internally consistent procedure in 2011, a method that also involved all countries, not just a small group. Our direct analysis of the ring countries, comparing the inflation rates from one ICP to the next, and comparing it

with inflation rates from national CPIs, suggested that the PPPs for the Western Asia, Africa, and Asia regions might have been inflated by about a quarter. A different approach, based on an analysis of the Balassa-Samuelson or Penn effect in 2005 and 2011, is consistent with a similar overestimation in 2005. This last method is conditional on assuming that 2011 is correct, and 2005 not, so it is less strong than the method that looks at inflation in the ring countries. But both sets of results are consistent with the up and down pattern shown by the discrepancies between actuals and interpolations in 2005 and 2011.

It is not our purpose in this paper to discuss how these results should be used. But our findings suggest that the ICP 2011 estimates are the most accurate that we have, and provide no grounds for doubting them. Indeed, there are other major improvements in 2011 that we have not discussed, including much more rigorous validation of data, and much more attention to the local currency national accounts supplied by the countries. If this is correct, then the revisions that need to be undertaken are to long-standing previous estimates, a process that is likely to be less than straightforward.

Given our results, one way to do correct 2005 would be to preserve the within-region PPPs for each of the regions in 2005, but to revise the regional linking. This could be done by back casting (backward extrapolation from 2011 to 2005) for each country and then aggregating GDP and consumption up to regions. The regional totals would then be allocated to the countries in proportion to the original 2005 estimates, which preserves the within-region totals. This is essentially the Country Aggregation with Redistribution (CAR) method that was used to link the regions in

2011. As far as the future is concerned, it would also be extremely desirable to put the program on a continuous basis, with results updated from year to year, which should avoid the large revisions that characterized the last two rounds.

6. Citations

Chandy, Lawrence, and Homi Kharas, 2014, "What do new price data mean for the goal of ending extreme poverty?" Brookings, May 5, <http://www.brookings.edu/blogs/up-front/posts/2014/05/05-data-extreme-poverty-chandy-kharas>

Deaton, Angus, 2001, "Counting the poor: problems and possible solutions," *World Bank Research Observer*, 16:2, 125–47.

Deaton, Angus, 2010, "Price indexes, inequality, and the measurement of world poverty," *American Economic Review*, 100:1, 5–34.

Deaton, Angus, 2012, "Calibrating measurement uncertainty in purchasing power parity exchange rates," incomplete paper, http://siteresources.worldbank.org/ICPINT/Resources/270056-1255977254560/6483625-1338834270350/ADeaton_CalibratinguncertaintyinPPP.pdf

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

Diewert, Erwin, 2013, "Methods of aggregation above the basic heading level: linking the regions," Chapter 6 in Fred Vogel and D. S. Prasada Rao, *Measuring the real size of the world economy: the framework, methodology and results of the international comparison program*, Washington, DC. The World Bank, pp. 169–96.

Dykstra, Sarah, Charles Kenny, and Justin Sandefur, 2014, "Global absolute poverty fell by almost half on Tuesday," Center for Global Development, May 2, <http://www.cgdev.org/blog/global-absolute-poverty-fell-almost-half-tuesday>

Inklaar, Robert, 2013, "Price levels and economic growth: making sense of revisions to data on real incomes: a comment," *Review of Income and Wealth*, 59:4, 614–22.

Kravis, Irving, Alan Heston, and Robert Summers, 1982, *World product and income: international comparisons of real gross product*, Baltimore, MD. Johns Hopkins University Press for the World Bank.

McCarthy, Paul, 2011, "Extrapolating PPPs and comparing ICP benchmark results," International Comparison Project,
http://siteresources.worldbank.org/ICPINT/Resources/270056-1255977254560/6483625-1291755426408/18_ICPBook_Extrapolating_FINAL.pdf

Obstfeld, Maurice and Kenneth Rogoff, 1996, *Foundations of international macroeconomics*, MIT Press.

Ravallion, Martin, 2012, "Price levels and economic growth: making sense of revisions to data on real incomes," *Review of Income and Wealth*, 59:4, 593–613

World Bank, 2014, *Summary of results and findings of the 2011 international comparisons program*, International Comparison Program,
<http://siteresources.worldbank.org/ICPINT/Resources/270056-1183395201801/Summary-of-Results-and-Findings-of-the-2011-International-Comparison-Program.pdf>

Table 1**Regional revisions from extrapolations in ICP 2005 and ICP 2011**

Region	Mean log ratio 2005	Mean log ratio 2011	s.d. log ratio 2005	s.d. log ratio 2011	Number of countries
Africa	0.57	-0.24	0.27	0.14	32
Asia & Pacific	0.46	-0.34	0.27	0.08	15
CIS	0.40	-0.15	0.24	0.25	5
Eurostat/OECD	0.08	-0.00	0.12	0.01	40
Latin America	0.26	-0.13	0.21	0.06	7
Western Asia	0.38	-0.57	0.16	0.11	6

Notes: CIS is the Confederation of Independent States. The means are calculated only for countries that are common to both rounds, and are unweighted country means of the logarithm of the ratio of the ICP consumption PPP to the previously extrapolated consumption PPP.

Table 2: Data on PPPs, exchange rates, and inflation from 2005 to 2011

Country	PPP 2005	PPP 2011	XR 2005	XR 2011	Relative Inflation rate
Brazil	2.801	2.720	4.418	2.679	0.114
Chile	684.1	641.0	1018.4	780.8	-0.132
Cameroon	649.8	404.2	959.0	756.2	0.003
Egypt	3.570	3.275	10.51	9.474	0.465
Estonia	15.51	13.64	22.87	18.03	0.104
UK	1	1	1	1	0
Hong Kong	9.764	9.128	14.15	12.47	-0.017
Jordan	0.6786	0.533	1.291	1.138	0.158
Japan	218.2	160.37	200.4	127.9	-0.186
Kenya	64.80	61.77	137.4	142.3	0.540
Sri Lanka	80.18	75.06	182.7	177.2	0.426
Malaysia	3.675	2.479	6.891	4.904	-0.016
Oman	0.3393	0.292	0.6909	0.617	0.134
Philippines	39.72	31.49	100.2	69.41	0.106
Senegal	604.3	439.9	959.0	756.2	-0.009
Slovenia	287.2	252.4	350.3	276.1	-0.019
South Africa	7.735	8.132	11.56	11.64	0.206
Zambia	5235.3	4111.0	8115.5	7789.5	0.395

Notes: PPPs are calculated using consumption data from ring countries only; the aggregation uses a two-stage CPD and weighted-CPD method below and at BH level. The UK is the base country, and all PPPs and exchange rates take GBP as numeraire. The inflation rate is the change in the log of each country's CPI from the World Development Indicators minus the change in the log of the British CPI, so that the column shows the inflation rate relative to Britain. The exchange rates and PPPs for Estonia and Slovenia in 2011 have been multiplied by 15.6466 and 239.64 respectively, which are the rates at which each country entered the Eurozone. If extrapolation of PPPs worked perfectly, the log of the second column minus the log of the third column should be equal to the last column.

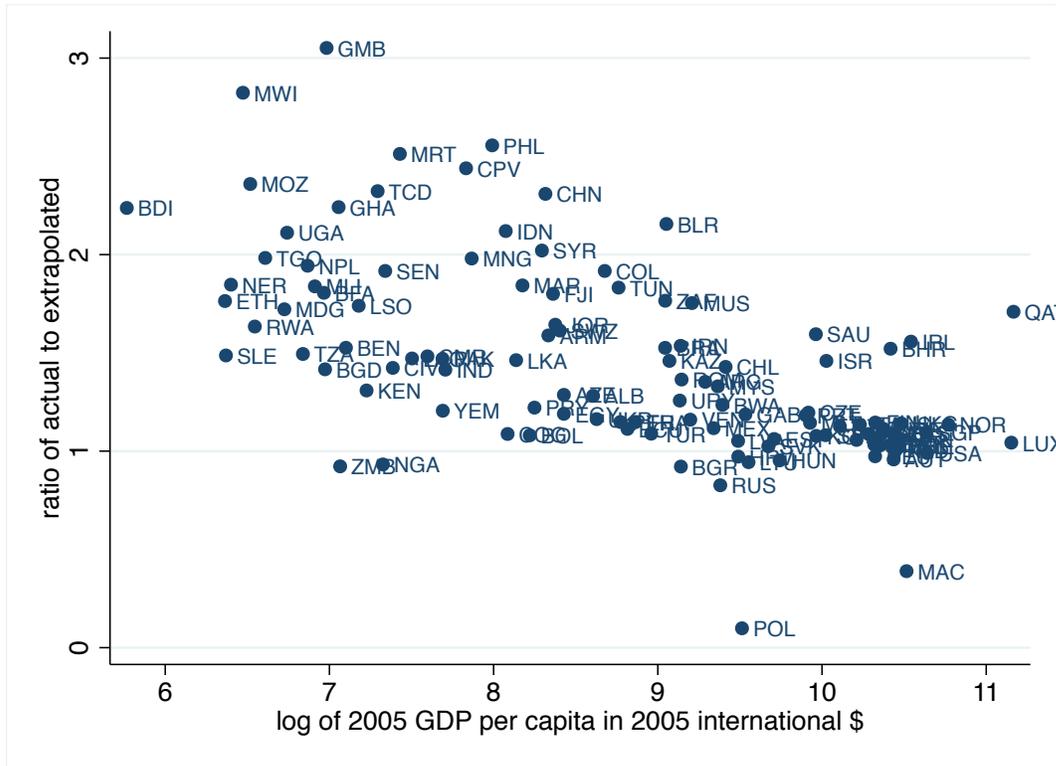


Figure 1: Ratio of ICP to extrapolated consumption PPPs, ICP 2005

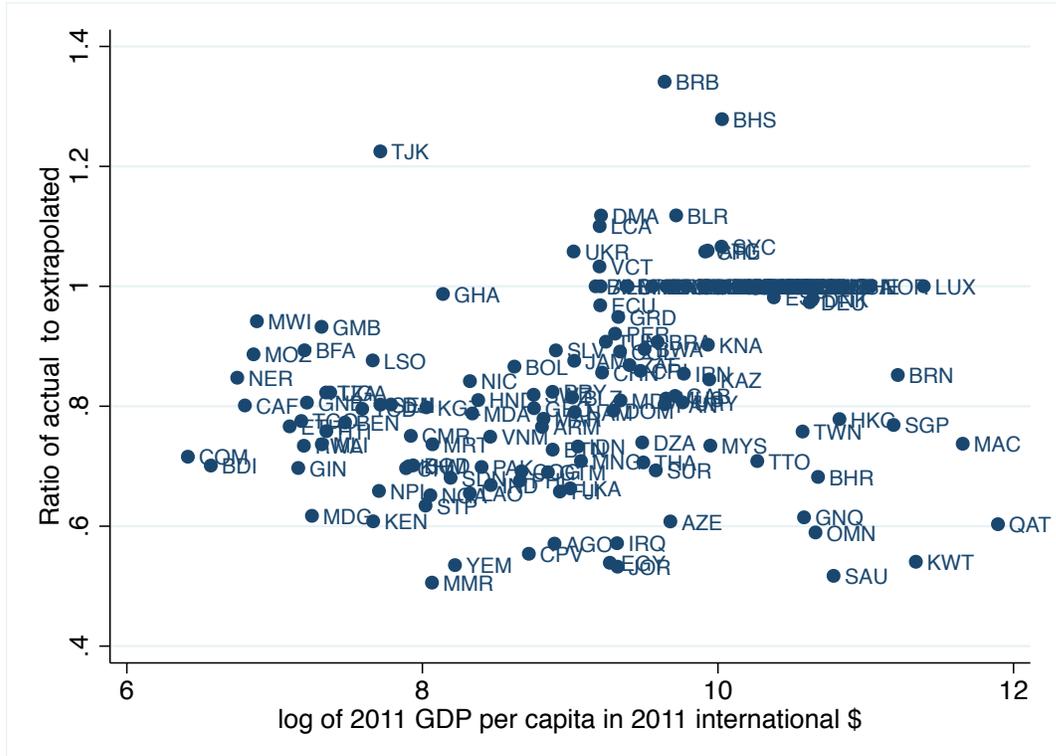


Figure 2: Ratio of ICP to extrapolated consumption PPPs, ICP 2011

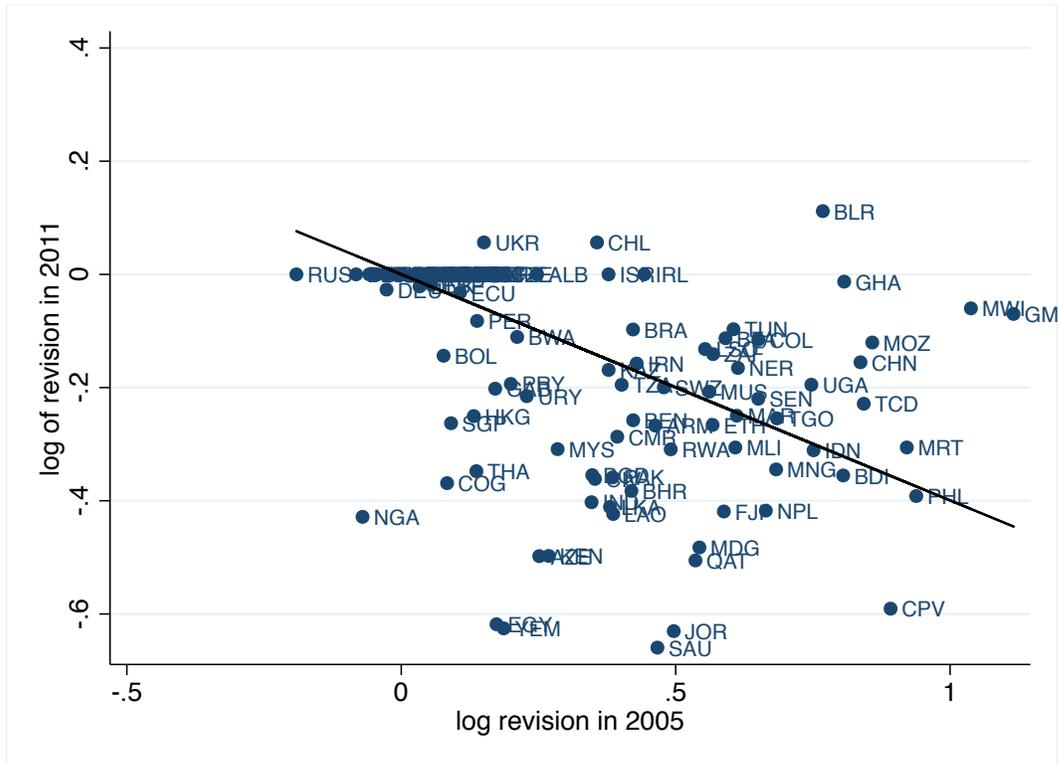


Figure 3: Logs of revision ratios in 2005 and 2011

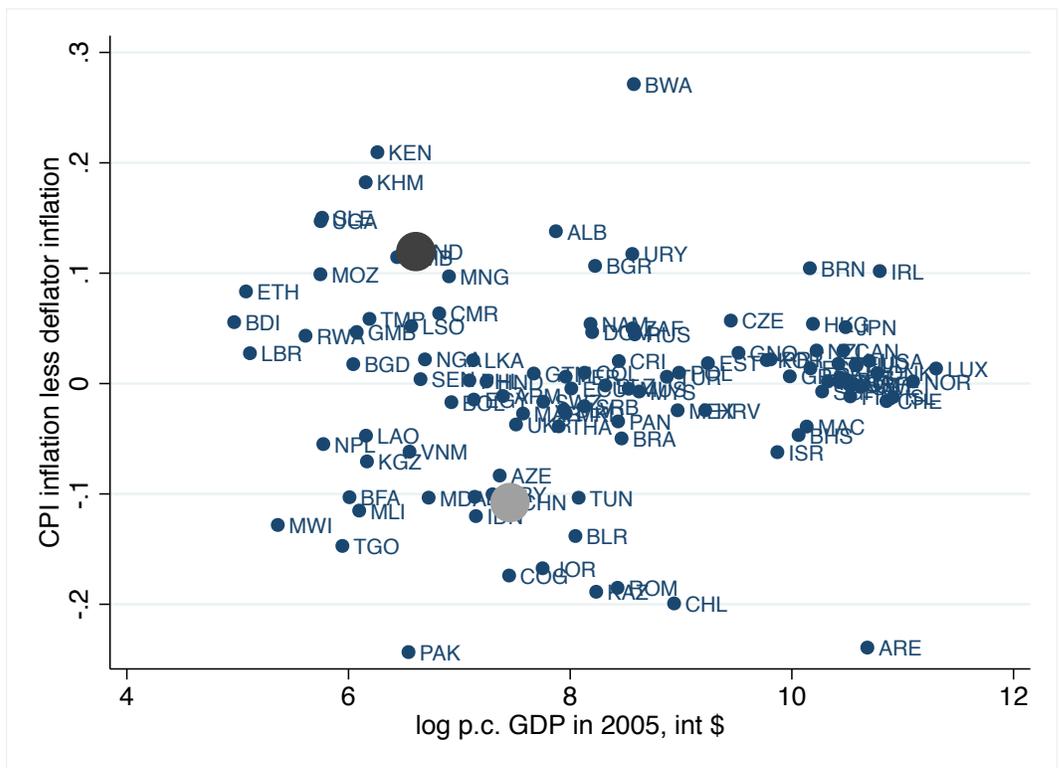


Figure 4: CPI inflation minus consumption deflator inflation 2005-2011

India (dark grey) and China (light grey) enlarged

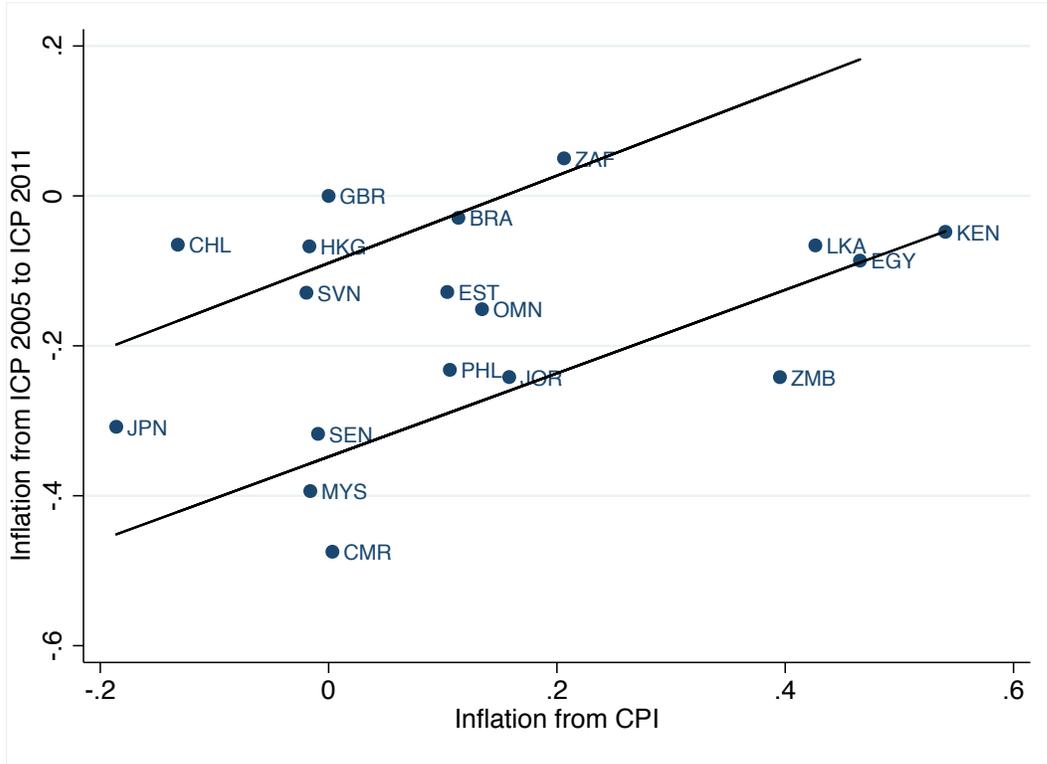


Figure 5: ICP based inflation versus CPI inflation, 2005 ring countries, 2005 to 2011

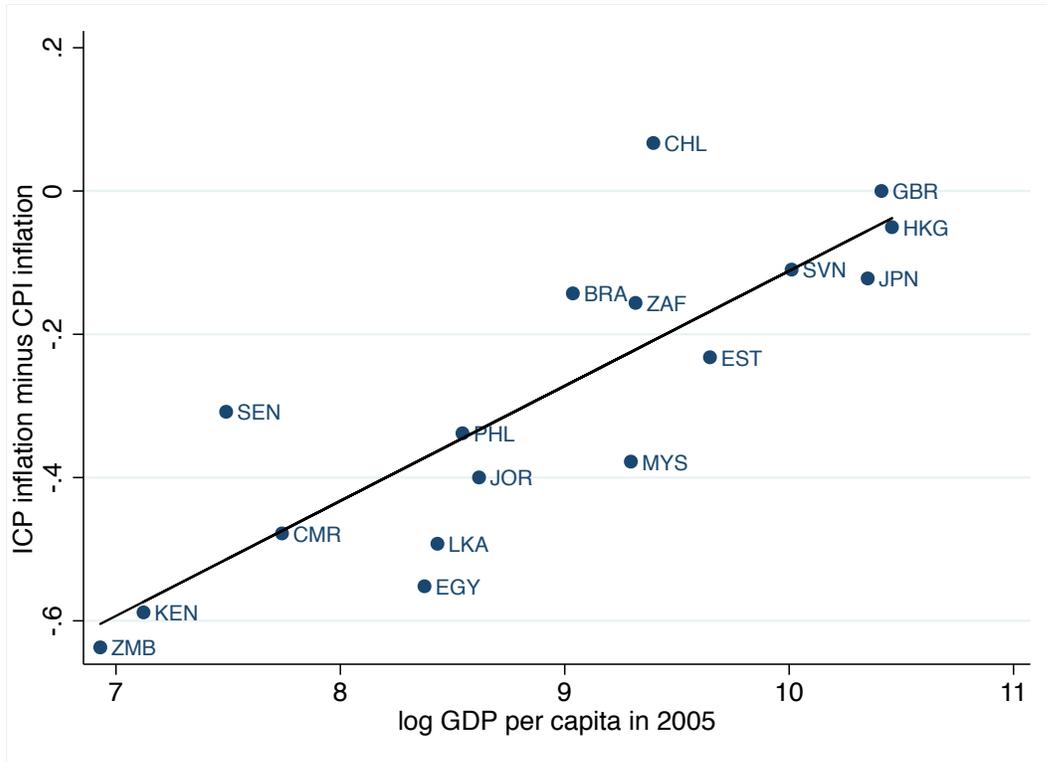


Figure 6: ICP minus CPI inflation and GDP per capita

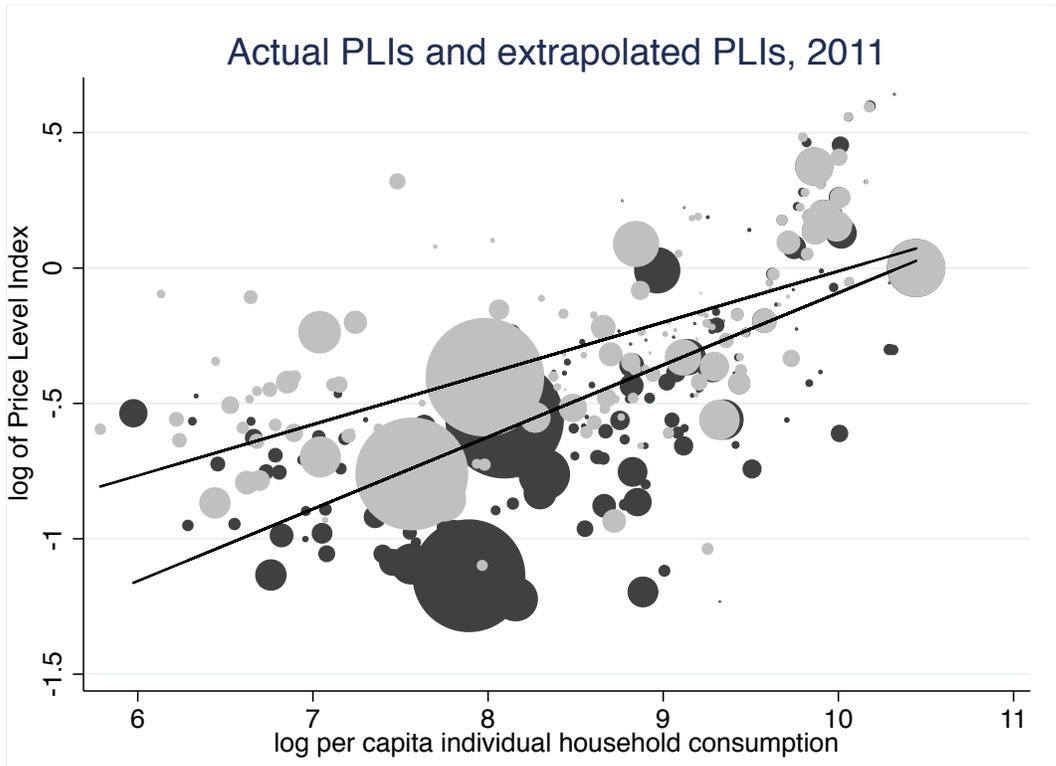


Figure 7: Extrapolated (light gray) and actual (dark grey) PLIs, 2011

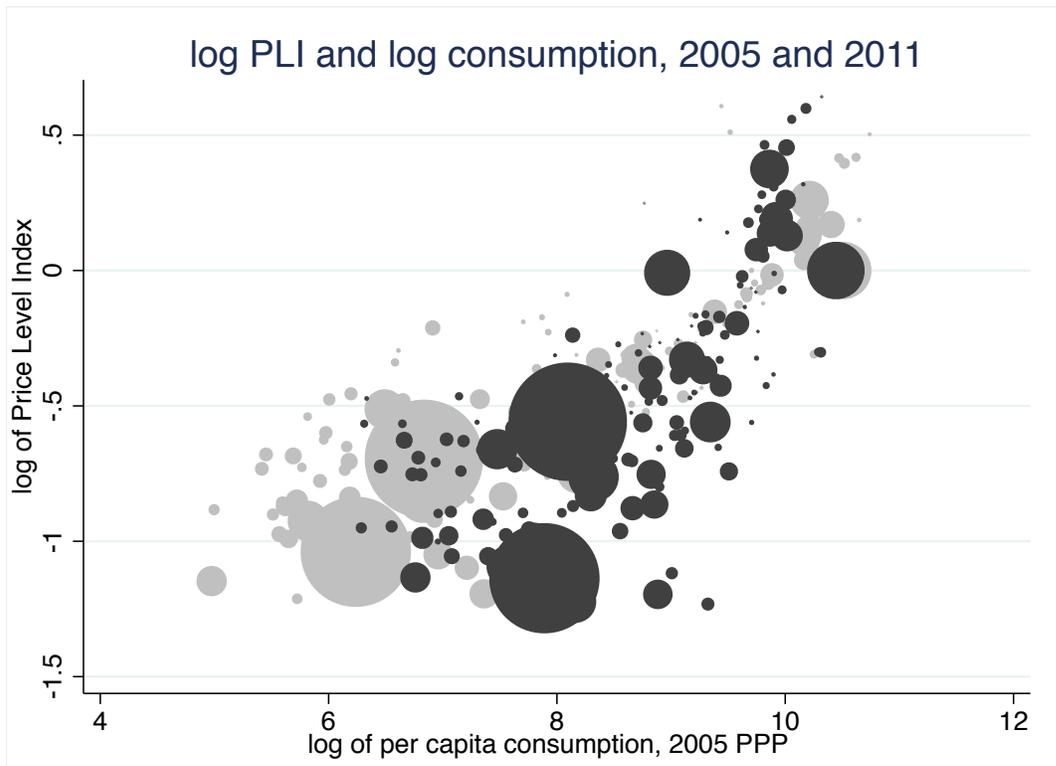


Figure 8: PLIs and consumption, 2005 (light) and 2011 (dark) ICPs