Revisions to Global Income Comparisons – the Case of ICP 2011

Robert Inklaar (University of Groningen, Netherlands)

Paper Prepared for the IARIW 33rd General Conference

Rotterdam, the Netherlands, August 24-30, 2014

Session 6A

Time: Thursday, August 28, Afternoon

Revisions to global income comparisons – the case of ICP 2011

Robert Inklaar*

Groningen Growth and Development Centre, University of Groningen

Abstract: The World Bank recently released data on relative price and income levels for 2011 based on a global survey of prices. This led to an upwards revision of income levels by about 25 percent for the average country relative to the US. The earlier estimates were based on a 2005 global survey of prices, raising the question which is to blame: the previous survey, the current survey or the extrapolation method that is used for updating relative price and income levels. Here I argue that the current survey is methodologically superior to the previous one and that no extrapolation method can resolve the inconsistency. I propose a revision to the previous survey results that does restore consistency.

Acknowledgements: the author would like to thank Angus Deaton for helpful comments.

Correspondence to: Robert Inklaar, University of Groningen, Faculty of Economics and Business, PO Box 800, 9700 AV Groningen, The Netherlands (R.C.Inklaar@rug.nl).

Introduction

The World Bank (2014) recently published data on comparative prices and expenditure levels for the year 2011 as part of its International Comparison Program (ICP). These data will serve as the official yard stick for comparing the size of economies and gauging the level of global poverty for years to come, making this data release of great importance.[1] It is further an undoubted milestone: in comprehensively covering 177 countries representing 97 percent of the world population, this is a truly global comparison. Most important for our purposes, though, is that ICP 2011 is the second time there has been a global comparison of this scale and quality, with the first held in 2005.[2] With two sets of relative prices across countries (typically referred to as purchasing power parities, PPPs) the question is to what extent these are consistent with (independent) information on inflation, i.e. changes in prices over time. A lack of consistency would mean that at least one of the pieces of information has to be downplayed or discarded.[3] This is of particular relevance for datasets that provide data on relative income levels over time, like the Penn World Table (Feenstra, Inklaar and Timmer, 2013).[4]

In the first systematic comparison of results from ICP 2005 and ICP 2011, Deaton and Aten (2014) show that the two comparisons are far from consistent.[5] This could be due to two reasons:

1. Relative inflation rates are not appropriate for predicting the change in PPPs.

---

[1] See Deaton and Heston (2010), Chen and Ravallion (2010) and Deaton (2010) for broader discussions on the ICP data and their context in development research and policy.

[2] There have been five earlier comparisons, in 1970, 1975, 1980, 1985 and 1996, but these were much less consistent in their data collection and methodologies.

[3] See e.g. Hill (2004) and Rao, Rambaldi and Doran (2010).

[4] And this is a topic that has attracted numerous earlier analyses, including Krijnse Locker and Faerber (1984), Summers and Heston (1991, 1993), Dalgaard and Sørensen (2002), Ravallion (2013) and Inklaar (2013).

[5] Though this point has also been noted by Chandy and Kharas (2014), Dykstra, Kenny and Sandefur (2014) and Ravallion (2014).

2. The ICP 2005 PPPs were flawed.

The World Bank (2014) seems to argue for reason 1, pointing to the conceptual reasons why inflation rates may fail to adequately capture PPP changes and reason 2, when they discuss the substantial improvements in the computation of PPPs in ICP 2011 compared with ICP 2005.[6] But while their arguments may be relevant in principle, the World Bank (2014) does not give a quantitative accounting or reconciliation. Deaton and Aten (2014) give greater weight to reasons 2, arguing that part of the ICP 2005 methodology – the comparison across the ICP regions – lead to a systematic bias. However, their analysis is limited to household consumption, does not provide a detailed quantitative assessment of reason 1 and does not implement an alternative to ICP 2005 that would not suffer from the systematic bias they identify. Flaws in ICP 2011 have not been explicitly considered as a separate reason, mostly because methodological flaws that might bias ICP 2011 have been present in ICP 2005 as well, while some of the concerns about ICP 2005 have subsequently been addressed in ICP 2011.[7]

The aim of this paper is to systematically and quantitatively analyse the two reasons. The starting point is to compare the PPPs from ICP 2011 with those predicted from ICP 2005 using relative inflation. Based on this, I propose a set of statistics that can be used to determine how well or how poorly a particular prediction fares. I will then consider two alternative approaches to predicting changes in PPPs, taking ICP 2005 and ICP 2011 as given. The first approach uses the systematic relationship between price and income levels, similar in spirit to Ravallion (2013); the second approach uses inflation information but at a more detailed level. Neither approach proves satisfactory: there are still important systematic biases in the predicted PPPs. In the next step, I will take inflation extrapolation and ICP 2011 as given but consider an alternative to ICP 2005. Following a procedure proposed by Deaton and Aten (2014), I show that

[6] See McCarthy (2013).

[7] See World Bank (2013) and (2014, p25-27). Notable improvements were made in accounting for the representativeness of individual products, estimating hard-to-measure prices such as for rents and government services, and linking relative prices across regions.

predictions are much improved if ICP regions are linked in a similar way in 2005 as in 2011. This procedure is based, in part, on results from ICP 2011 but it leaves the most reliable element of ICP 2005, the within-region relative prices, untouched. Compared with the original ICP 2005, the adjusted ICP 2005 shows a very similar relationship between prices and income levels but between-country income inequality is considerably lower. In effect, this shifts some of the observed reduction in income inequality from the 2005-2011 period to the pre-2005 years (see Milanovic, 2012).

Initial comparison

The initial step in the analysis is to compare the new ICP 2011 figures on GDP per capita and household consumption per capita with the corresponding figures extrapolated from ICP 2005. For ICP 2011, I use the results as published in the Summary Report (World Bank, 2014) and specifically the PPPs and expenditure levels for GDP and final household consumption expenditure ('consumption'). For ICP 2005, I use the PPPs as originally published in World Bank (2008).[8] The combined dataset is constrained in its country coverage mostly by the coverage of ICP 2005, which spanned 146 countries. Argentina, Syria and Lebanon participated in ICP 2005 but not in ICP 2011 and Zimbabwe participated in ICP 2005 but the coinciding period of hyperinflation left the results so unreliable that the World Bank does not publish a full set of PPP and exchange rate results. That leaves us with a dataset of 142 countries.

Let P^{τ}_{Yjt} be the PPP for GDP (Y) in country j at time t based on a benchmark year τ and let p_{Yjt} be the GDP deflator in year t and country j. Whenever $t \neq \tau$ the PPP is extrapolated from the benchmark year as:

$$(1) \qquad P^{\tau}_{Yjt} = P^{\tau}_{Yj\tau} \times \frac{p_{Yjt} / p_{Yj\tau}}{p_{Ybt} / p_{Yb\tau}}$$

[8] Or, to be more precise, the ratio of the PPP over the exchange rate. This avoids having to explicitly control for changes in currency, such as the adoption of the euro.

The PPPs are defined relative to a base country *b* – usually the United States – and PPPs are extrapolated to years other than the benchmark year using relative inflation rates, the same procedure as followed for the World Development Indicators.[9] For the purpose of poverty calculations the emphasis is typically on household consumption *C*, in which case extrapolation is done using the consumer price index (CPI) rather than the GDP deflator.[10]

We can then compare GDP per capita (*y*) or household consumption per capita (*c*) based on the PPP extrapolated from ICP 2005 with the same numbers based on the new ICP 2011 benchmark:

$$(2) \qquad d_{Yj2011} = \frac{y_{j2011} / P^{2011}_{Yj2011}}{y_{j2011} / P^{2005}_{Yj2011}} - 1 = \frac{P^{2005}_{Yj2011}}{P^{2011}_{Yj2011}} - 1$$

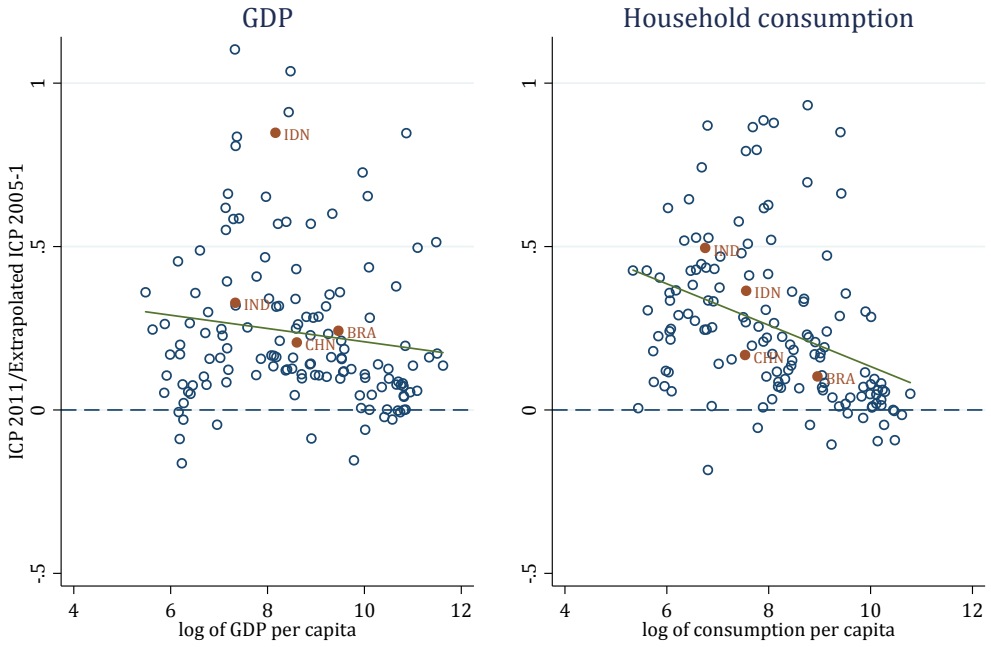
As all subsequent comparisons are for the year 2011, we drop the subscript *t*. Figure 1, show a plot of d_y against the log of GDP per capita converted using exchange rates *E*, $\log\left(y_j / E_j\right)$ and d_c against consumption per capita, $\log\left(c_j / E_j\right)$. The GDP and consumption per capita numbers are those reported in World Bank (2014). The figure illustrates the very large revisions to PPPs that came with ICP 2011. As a result of the PPP revisions, by far most countries show upward revisions to GDP and consumption per capita levels. As the range of the graphs shows, relative income and consumption levels have increased by more than half in a sizeable number of countries. Furthermore, there is a clear negative correlation with the highest-income countries typically only showing small

[9] Note, though, that the World Development Indicators also take into account the more frequent benchmark PPP estimates for countries in the OECD and European Union. This is also the case for PWT, which in addition extrapolates PPPs at a more detailed level than overall GDP, see Feenstra et al. (2013).

[10] An alternative would be the implicit price deflator of household consumption expenditure from the National Accounts but the following results are very similar.

revisions. The lower-income countries typically show larger revisions and this effect is strongest for consumption.[11]

Figure 1, Differences between extrapolating ICP 2005 and ICP 2011



Note: GDP and consumption per capita are in exchange-rate converted US$.
Extrapolated from ICP 2005 using GDP/consumption inflation rate of each country relative to the US.

To summarise the patterns of Figure 1 and compare these to alternative estimates in the remainder of this paper, we use three statistics, namely the average difference across the set of *J* countries, $\bar{d}_i = \frac{1}{J}\left(\frac{P_{Yj2011}^{2005}}{P_{Yj2011}^{2011}} - 1\right)$,[12] the average squared difference, $\bar{d}_i^2 = \frac{1}{J}\left(\frac{P_{Yj2011}^{2005}}{P_{Yj2011}^{2011}} - 1\right)^2$, and the slope coefficient of the regression $d_{ij} = a_i + b_i \log\left(e_{ij}/E_j\right)$, all for $i \in Y, C$. These summary measures are best understood by viewing the problem as a forecasting exercise: until the release of the benchmark ICP 2011 PPPs, the typical approach was to forecast the

[11] The difference between the GDP and consumption pattern suggests that revisions to the price levels of investment and government consumption are not related with income levels.

[12] The median difference does not show a qualitatively different pattern.

PPPs for 2011 using equation (1).[13] Statistic $\overline{d}_i$ is then a measure of the forecasting bias, $\overline{d}_i^2$ a measure of the forecasting uncertainty[14] and b_i an indication whether the bias varies systematically with GDP or consumption per capita.[15]

Given the challenge of comparing prices of more than a thousand individual products across almost 200 countries, it is not surprising that measurement error is present, which would lead to $\overline{d}_i^2$ being larger than zero.[16] Indeed, this can be a rationale for conducting regular benchmark comparisons: since the PPP of a country in a given year will be measured with error, it is good to have measurements in multiple years to avoid relying too heavily on a single (error-prone) observation.[17] That said, $\overline{d}_i^2$ is still a helpful measure since a method of extrapolating PPPs that would lead to a lower value of $\overline{d}_i^2$ would be preferable. The average bias $\overline{d}_i$ and the systematic variation in the bias measured by b_i are arguably the most worrisome measure as they imply a shift in the world income distribution. This is most obvious for b_i, as this measures whether lower-income

[13] An exception was Ravallion (2013), who has argued for a so-called 'dynamic Penn effect'. See also the discussion below.

[14] In the terminology of the forecasting literature, this would be the mean squared prediction error, MSPE, see West (2006).

[15] Note that the average difference and average squared difference are not independent of the numeraire country. However, since typical comparisons are almost exclusively made with the US as the numeraire this is not a major drawback. However, see Diewert (2009) for alternative, symmetric measures.

[16] See Deaton and Heston (2010) for a good discussion of the conceptual and practical challenges of cross-country price measurement. Note also that measurement error may not just be a problem for measuring PPPs but also for tracking inflation. The errors in inflation measurement should typically be smaller since it is more likely that a particular product can be priced from one period to the next than in two (potentially disparate) countries, though if information on spending patterns are not regularly updated, measured inflation could well start to substantially deviate from true inflation.

[17] See also e.g. Hajargasht and Rao (2010) on the estimation of PPP standard errors from the variation of individual product prices around the overall GDP or consumption PPP.

countries show systematically larger (or smaller) differences, but because all PPPs are given relative to the United States, a non-zero b_i implies an average shift in prices – and thus expenditure levels – relative to the US.

Table 1, Differences between ICP 2011 and extrapolation from ICP 2005 – summary statistics

	GDP	Consumption
Average difference	0.237***	0.254***
Average squared difference	0.113	0.122
Coefficient on log(expenditure/capita)	–0.020*	–0.063***
	(0.012)	(0.011)

Notes: summary statistics based on differences in Figure 1 for 142 countries. Robust standard error of the regression coefficients shown in parentheses below the coefficients. * denotes a variable significantly different from zero at a 10%-level, ** at 5%-level, *** at a 1%-level.

Table 1 shows the summary statistics corresponding to the two panels in Figure 1. The average country has expenditure levels in ICP 2011 that are about a quarter higher than based on extrapolations from ICP 2005. Put differently, the average country is a quarter richer relative to the US than estimated previously, implying a sizeable decline in cross-country income inequality. The average squared difference is also comparable between the GDP and consumption, at 11 and 12 percent. The main difference between the GDP and consumption extrapolation errors is that the coefficient on expenditure per capita is much larger for consumption than for GDP, though both are significant.

In the remainder of this paper, the statistics in Table 1 will serve as a baseline to help determine whether a particular alternative approach is helpful in improving the consistency between ICP 2005 and ICP 2011. In the next section, I consider three alternative approaches to using ICP 2005 information to estimate 2011 PPPs, so centring on reason 1 from the introduction. The subsequent section focuses on alternatives for the ICP 2005 PPPs, so reason 2.

Alternative extrapolation approaches

Using relative inflation rates to extrapolate PPPs, as in equation (1), is the typical approach and can easily be motivated given that PPPs measure the relative level

of all prices in the economy[18] and inflation measures the rate of change in those same prices. In this section, we consider two alternatives to this simple approach. The first alternative was suggested by Ravallion (2013) and exploits the correlation between relative prices and income levels, what has been referred to as the Balassa-Samuelson effect or the Penn effect.[19] The second alternative addresses the problem that expenditure shares of individual products are incorporated differently in the computation of PPPs than in the computation of national inflation rates, a problem that is most lucidly expressed in Deaton and Aten (2014).

Balassa-Samuelson extrapolation

Since the first systematic estimates of PPPs, it has been recognised that the ratio of the PPP over the exchange rate is typically higher in richer countries. In regression form this means there will be a positive β in the following equation:

$$\log\left(P_{Yj}/E_j\right) = \alpha + \beta \log\left(y_j/E_j\right) + \varepsilon_j \quad (3)$$

Balassa (1964) and Samuelson (1964) explained this positive relationship was from the difference between the traded sector and the non-traded sector of the economy. The traded sector would have very similar prices across countries and be more amenable to productivity improvements than the non-traded sector. Economic development – from increased productivity in the traded sector – leads to higher economy-wide wages. This in turn leads to pressure on prices in the non-traded sector and thus increases in the overall (average) price level of the economy.

Ravallion (2013) proposed to use the regularity from equation (3) to predict changes in PPPs between one benchmark round and the next. In what refers to to as the 'Dynamic Penn effect', he explains changes in the PPP over exchange rate ratio from changes in GDP per capita (see his model in equation (4) below). He concluded that accounting for economic growth would reduce the need for large

[18] With the exception, in the case of ICP, of prices of exports and imports; see the next section.

[19] See Balassa (1964) and Samuelson (1964) for the famous exposition, Samuelson (1994) for coining the term 'Penn effect' and Feenstra et al. (2013) for a recent analysis in this context.

data revisions. In my view there are two problems with the approach of Ravallion (2013). The first problem, discussed in Inklaar (2013), is that his proposed test does not isolate the effect of economic growth on changes in PPPs but also includes the effect of inflation on PPPs (i.e. the standard extrapolation) and the effect of exchange rates. A properly specified test shows no systematic relationship between PPP changes and economic growth.

The second problem, which I aim to address here, is that Ravallion (2013) used 'ex-post prediction': he explained the changes in PPPs from ICP 1993 to ICP 2005 and claimed that this gave superior results for 2005 compared with inflation-based extrapolation between 1993 and 2005 as in equation (1). However, this does not address the true issue, namely how to predict new PPPs if the actual PPP data is not yet available.

I therefore consider three models that could be used for out-of-sample prediction. The first model I estimate is equation (3), so using the 2005 relationship and 2011 expenditure/capita data to predict 2011 PPP over exchange ratios. The second model is based on the work of Hassan (2012), who argues that the price-income relationship is non-linear rather than linear. The theoretical argument is that in the initial phase of development, productivity growth is typically concentrated in agriculture, which at that point is a mostly non-traded sector. When manufacturing starts to represent a sizeable fraction of the economy, the traditional Balassa-Samuelson argument starts to hold. Like in Hassan (2012), this argument is tested by adding squared GDP/capita to equation (3).

The third model is the 'encompassing' model proposed by Ravallion (2013), which explains changes in the PPP over exchange rate ratio using changes in GDP per capita, relative inflation and changes in the exchange rate:

$$(4) \qquad \Delta \log\left(P_{Yj}/E_j\right) = \alpha + \beta_1 \Delta \log\left(y_j/E_j\right) + \beta_2 \Delta \log\left(p_{Yj}/p_{Yb}\right) + \beta_3 \Delta \log\left(E_j\right) + \varepsilon_j ,$$

where the Δ-operator indicates the difference between two periods. The coefficients in equation (4) have no clear (theoretical) interpretation because, first, relative inflation and exchange rate changes tend to be highly correlated. So

the effect of relative inflation is estimates *conditional* on the effect changes in the exchange rates. Second, the change in GDP per capita compares the nominal, exchange-rate converted GDP per capita level in two periods and is thus affected by inflation and exchange rate movements.[20] But even if the coefficients cannot readily be individually interpreted, it could yield a helpful prediction of changes in PPPs. If that were the case, though, a more thorough analysis would be needed to establish whether the coefficients are stable over time. One reason why they may not be is that coefficient estimates based on only a single change over time (i.e. 1993-2005) would be sensitive to shocks during this period that are correlated with the explanatory variables.[21]

It is important to realise that using these models to form a prediction implicitly assumes that the original inconsistency between ICP 2005, ICP 2011 and relative inflation between 2005 and 2011 is due to problematic inflation estimates. The first two, 'static' models avoid using inflation all together, relying fully on the level of nominal expenditure per capita converted to US dollars using market exchange rates The third, 'dynamic' model does use relative inflation, but uses econometric estimates of its importance in the (joint) determination of prior changes to give it an 'adjusted' weight. To the extent that the Deaton and Aten (2014) hypothesis of flaws in ICP 2005 can be used to construct a more comparable version of that benchmark, the subsequent regression analysis may need to be revisited.

To estimate the two 'static' Balassa-Samuelson models (cf. equation (3)), I supplement the 2005 PPP and exchange rate data from World Bank (2008) by data on GDP per capita drawn from the December 2013 version of the UN National Accounts Main Aggregates Database. This ensures that all National Accounts revisions since 2008 are taken into account and this data covers the full set of 142 countries. For the 'dynamic' Balassa-Samuelson model (cf. equation

[20] See Inklaar (2013) for a more detailed discussion.

[21] An example from the more recent period would be the 2008 financial crisis and subsequent Great Recession that has had a more substantial effect on higher-income countries.

(4)), I use PPPs for 1996,[22] and UN National Accounts data on GDP per capita, inflation and exchange rates. This second dataset is limited to 96 countries.

Table 2, Predicting PPPs and changes in PPPs from the Balassa-Samuelson relationship

	Log of price level, 2005				Change in price level, 1996-2005	
	GDP	Consumption	GDP	Consumption	GDP	Consumption
	(1)	(2)	(3)	(4)	(5)	(6)
log(expenditure/capita)	0.215***	0.211***	-0.504***	-0.569***		
	(0.0124)	(0.0132)	(0.112)	(0.130)		
log(expenditure/capita)^2			0.0440***	0.0507***		
			(0.00686)	(0.00830)		
Change in expenditure/capita					0.384**	0.0169
					(0.149)	(0.113)
Relative inflation					-0.0004	0.632***
					(0.308)	(0.202)
Change in exchange rate					0.0278	-0.502***
					(0.314)	(0.189)
Constant	-2.368***	-2.076***	0.445	0.803	-0.147**	0.00136
	(0.103)	(0.108)	(0.446)	(0.489)	(0.0741)	(0.0561)
Observations	142	142	142	142	96	96
R^2	0.716	0.661	0.788	0.740	0.241	0.221

Notes: columns (1) through (4) explain the log of the 2005 PPP over exchange rate ratio using the log of expenditure per capita (in 2005). Column (1) is identical to equation (3), while column (2) replaces the GDP PPP and GDP per capita by the household consumption expenditure PPP and expenditure per capita level; columns (3) and (4) add the square of expenditure per capita. Column (5) corresponds to equation (4) and column (6) replaces the GDP PPP, GDP per capita and inflation as measured by the GDP deflator by their household consumption expenditure equivalent. Robust standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2 shows the estimation results with each of the three models estimated using GDP PPPs and GDP per capita levels and using consumption PPPs and consumption per capita levels. Columns (1) and (2) confirm the standard result that countries with higher income or spending levels have higher PPPs relative to their exchange rate. Columns (3) and (4) confirm the result of Hassan (2012) that an inverted U-curve is a better fit for the data than the linear model of columns (1) and (2). Columns (5) and (6) run the Ravallion (2013)-style encompassing regression for the change in PPP/exchange rate ratio between 1996 and 2005. Those results are less clear-cut. Note first that the correlation between relative inflation and the change in exchange rates is 0.97, which means that the coefficient on each is hard to interpret in isolation. In the GDP column,

[22] This 1996 'benchmark' combines the ICP surveys for 1993 for a number of regions with the 1996 PPP for the OECD/Eurostat region, see World Bank (2008b).

(5), the change in expenditure per capita is significant with no joint effect from relative inflation and exchange rate changes while the reverse is the case in column (6).[23] But despite these differences, both specifications explain 20-25 percent of the variation in the dependent variables, so they can be used for prediction purposes.

The specifications from columns (1)-(4) –when combined with 2011 expenditure per capita levels – directly yield a prediction of the 2011 PPP/exchange rate ratio. The specifications from columns (5)-(6) yield a prediction in the change of this ratio between 2005 and 2011 when combined with data on changes in expenditure levels, exchange rates and relative inflation between these years. Combined with the 2005 PPP/exchange rate ratio, this yields at 2011 predicted ratio.[24] As in Table 1, the predicted price levels can be compared with the observed, ICP 2011 results and the summary statistics of those comparisons are shown in Table 3 together with the baseline results from Table 1.

Table 3 shows how all three models broadly improve upon the baseline inflation extrapolation approach: the average difference is (significantly) smaller (even negative when using the non-linear Balassa-Samuelson model) and the average squared difference is 17 to 68 percent smaller. The relationship between the differences and expenditure per capita provides a more mixed picture, with the negative relationship for consumption only disappearing in the non-linear Balassa-Samuelson model but in that model, poorer countries are systematically *poorer* in terms of GDP per capita. So if one were willing to take the ICP 2005 results as given (a topic we will return to below) it is possible to provide more accurate predictions of the 2011 PPPs than based on relative inflation. However, no single model is clearly superior and, given that the estimated relationship are all not structural, there is no guarantee that a particular model continues to describe the data well into the future. Indeed, given the earlier remarks that the

[23] A model that only includes the change in expenditure per capita is significant with a coefficient of similar magnitude in for GDP and consumption, but the subsequent results are not materially affected if that model were used instead of the model reported in Table 2.

[24] In all specifications the prediction for the US is different from 1, so this normalisation is imposed on the predictions.

dynamic Balassa-Samuelson relationship as estimated here and by Ravallion (2013) will be very sensitive to shocks, it seems probable that the estimated relationship is not a good guide to the future.

Table 3, ICP 2011 versus predicted PPPs using the Balassa-Samuelson relationship

	Baseline	B-S	B-S & squared	Dynamic B-S
GDP				
Average difference	0.237***	0.116***	-0.074***	0.095***
Average squared difference	0.113	0.094	0.065	0.037
Coefficient on log(GDP/capita)	-0.020*	0.016	0.048***	-0.009
	(0.012)	(0.017)	(0.015)	(0.009)
Consumption				
Average difference	0.254***	0.003	-0.242***	0.163***
Average squared difference	0.122	0.057	0.080	0.066
Coefficient on log(consumption/ capita)	-0.063***	-0.029**	0.004	-0.053***
	(0.011)	(0.014)	(0.008)	(0.009)

Notes: See notes of Table 1 for a general explanation of the statistics. Baseline figures correspond to Table 1, 'B-S' corresponds to the predicted PPPs based on the specifications in columns (1) and (2) of Table 2; 'B-S & squared' corresponds to columns (3) and (4) and 'Dynamic B-S' corresponds to columns (5) and (6); see the main text for details of the prediction approach.

Extrapolation at the basic-heading level

The second alternative extrapolation does not suffer from this problem as it directly deals with the main conceptual reason why national inflation rates will not be an accurate measure of changes in PPPs over time. For tracking national inflation, an appropriate measure would weigh each product's price change using national expenditure shares. However, for comparing prices across countries – as PPPs do – an appropriate measure combines expenditure shares from different countries. This issue is discussed more formally in Deaton and Aten (2014) and will lead to a systematic bias from predicting PPP changes using relative (overall)

inflation. The direction of this bias will depend on differences in expenditure shares and (average) inflation of specific products.

This is also a problem with a clear solution, namely to avoid extrapolating at an aggregate level, but do this at a detailed product level, referred to as the basic heading (BH) level in ICP. The detailed extrapolated relative prices can then be aggregated using the same approach and same set of expenditure shares as used for computing the 2011 PPPs.

There are approximately 130 BH categories for which relative prices and expenditure shares are available in both ICP 2005 and 2011. So ideally, specific relative inflation rates for each of the 130 categories would be available and would be used for extrapolation. Instead, though, the only data source on inflation by product covering most countries in the world is the ILO and they only distinguish a few major products – food, clothing, rent and utilities – in addition to providing an overall consumer price inflation measure. Presumably, though, the available set of products account for a sizeable part of cross-country variation in expenditure shares especially between lower-income countries – where food and clothing account for larger shares – and higher-income countries – where rent and utilities are comparatively more important. The product inflation rates do not cover all consumer spending categories and data coverage for individual countries is incomplete. In those cases, there is no alternative to using overall consumer price inflation. For investment and government consumption, National Accounts deflators are available and for net exports the change in exchange rate is used since ICP uses exchange rates to convert net exports to a common currency.

The individual basic heading relative prices are aggregated to PPPs for GDP and household consumption using a GEKS procedure that is also used in the ICP.[25] The procedure I follow does not precisely mimic ICP as they follow a two-stage aggregation procedure: first across countries within broad regions and then

[25] The GEKS procedure is based on Fisher indices, which are in turn based on a Laspeyres PPP (using expenditure shares of the base country) and a Paasche PPP (using expenditure shares of the comparison country). See Balk (2008) for a more detailed treatment of these and other index number methods.

across regions, see World Bank (2014) and Deaton and Aten (2014). However, experiments show that this difference in procedure is of second-order magnitude compared to the differences that are the focus of this paper.

Table 4, ICP 2011 versus PPPs extrapolated at the basic heading (BH) level

	GDP		Consumption	
	Baseline	BH extrapolation	Baseline	BH extrapolation
Average difference	0.237***	0.194***	0.254***	0.223***
Average squared difference	0.113	0.097	0.122	0.128
Coefficient on log(expenditure/capita)	-0.020*	-0.014	-0.063***	-0.065***
	(0.012)	(0.011)	(0.011)	(0.012)

Notes: See notes of Table 1 for a general explanation of the statistics. Baseline figures correspond to Table 1, BH extrapolation refers to PPPs computed after extrapolating basic heading level PPPs from ICP 2005 using detailed inflation series and the aggregating using a GEKS procedure and 2011 expenditure shares, see main text for further details.

Table 4 presents the results from comparing the actual ICP 2011 PPP to the PPPs extrapolated from 2005 at the BH level; the baseline results from Table 1 are included for reference. The table shows that BH-level extrapolation modestly improves the statistics for overall GDP, but there are no notable changes for consumption. This suggests that the most substantive differences in expenditure shares and average inflation rates are at the top level of aggregation – between household consumption, investment, government consumption and nest exports – rather than between individual consumption categories. That would be good news for the Penn World Table, which takes such top level changes into account. However, even for GDP the average difference is still worryingly large and the bias for consumption is still systematically related with income level.

In summary, it does not seem possible to tweak the standard inflation extrapolation approach in such a way that ICP 2005 and 2011 are consistent with each other. An econometric approach, so eschewing the mechanical adjustment for relative inflation, could be more helpful but involves selecting the appropriate regression model and even when only focusing on the outcomes (rather than the precise theoretical underpinning), there is no clear favourite.

Alternative 2005 benchmark

The results from the previous section suggest that the ICP 2005 and 2011 comparisons differ in a systematic fashion because is a large average difference between the observed 2011 PPPs and those extrapolated from 2005 and this difference is systematically related with income levels.

One prominent candidate for explaining this systematic difference was advanced by Deaton and Aten (2014) and relates to how the regional comparisons are linked into a global comparison. In both ICP 2005 and ICP 2011, each region used its own list with product definitions as the basis for each of the basic heading level PPPs. To be able to link these price comparisons across regions, at least some countries must also compare a global list of comparable products. In ICP 2005, only a selection of 18 countries priced this global list; these are referred to as 'ring countries' (World Bank, 2008). However, the selection of ring countries is not innocuous as their price patterns may be systematically different from the average regional price pattern. Therefore, all countries priced the global product list (in addition to the regional list) in ICP 2011. Deaton and Aten (2014) compare PPP changes to relative inflation for the ring countries and they compare the Balassa-Samuelson relationship for the most-affected regions and find in both cases that in Africa, Asia and Western Asia, PPP in ICP 2005 may have been overstated by 20-25 percent. In this section, I will explore this possibility in more detail.

Within-region extrapolation

Deaton and Aten (2014) argue that the main reason for the differences between the extrapolated ICP 2005 and the observed ICP 2011 are the way in which the various ICP regions were linked together in ICP 2005. If this is indeed the case, we would expect that extrapolating within each region would lead to substantially smaller differences than extrapolating in the full global sample. In terms of equation (1), this means that in each region a base country b is selected and all PPPs and relative inflation rates are computed with that base country as

the numeraire. The results are invariant to which country is chosen as the base, so here I follow the ICP base country choice.[26]

Table 5, ICP 2011 versus PPPs extrapolated within each ICP region

	GDP				Consumption			
	$\overline{d}_Y$	$\overline{d}_Y^2$	b_Y		$\overline{d}_C$	$\overline{d}_C^2$	b_C	
Global	0.237***	0.113	-0.020*	(0.012)	0.254***	0.122	-0.063***	(0.011)
Africa	-0.412***	0.184	0.017	(0.011)	-0.127***	0.036	0.031	(0.026)
Asia	0.368***	0.175	-0.042	(0.030)	0.074***	0.013	-0.039***	(0.009)
CIS	-0.049	0.019	0.104***	(0.031)	-0.035	0.034	0.090	(0.086)
Eurostat/ OECD	0.003	0.006	-0.032**	(0.012)	-0.005	0.006	-0.055***	(0.015)
Latin America	0.066**	0.011	0.071*	(0.041)	0.068**	0.012	0.041	(0.044)
Western Asia	-0.002	0.012	-0.034	(0.031)	0.043	0.009	-0.041	(0.025)

Notes: the row 'Global' contains the summary statistics from Table 1; the columns contain the average difference $\overline{d}_i$, the average squared difference $\overline{d}_i^2$ and the slope coefficient of the differences on expenditure per capita, b_i for *i* equal to GDP, *Y*, or consumption, *C*. CIS stands for Commonwealth of Independent States. Extrapolation is done within each region, see the main text for discussion. See notes to Table 1 for further details and see World Bank (2014) for the country composition of each region.

Table 5 shows that the differences within regions are notably smaller than for the full global comparison. This is most prominently the case in Western Asia, where the average differences and the regression coefficients on expenditure per capita are not significantly different from zero and the average squared difference is around 1 percent. Large improvements are also seen in the Commonwealth of Independent States (CIS) and the Eurostat/OECD regions, with average differences close to zero and much smaller average squared differences.[27] The

[26] Africa – Nigeria, Asia – Hong Kong, CIS – Russia, Eurostat/OECD – Germany, Latin America – Peru, and Western Asia – Oman.

[27] In the Eurostat/OECD region, the revisions upon publication of ICP 2011 were even smaller than suggested here since Eurostat publishes annual benchmark PPP estimates and the non-EU OECD countries every three years, so 2008 was the most recent benchmark before 2011.

results for Africa and Asia, though, show that within-region extrapolation does not resolve all extrapolation differences, especially not for the extrapolation of GDP PPPs.[28] Still, compared to the results of the global extrapolation in the top row, the within-region extrapolation is much less problematic.[29]

Relinking the regions

The results in Table 5 provide further support for the Deaton and Aten (2014) argument that the approach to linking the regional results in ICP 2005 may be to blame for some of the large extrapolation differences. So what would be required is method that leaves the within-region PPP of ICP 2005 intact but uses a different approach for comparing across regions. Deaton and Aten (2014) propose the following procedure:

1. Extrapolate the 2011 PPPs backwards to 2005 at a global level, $P_{ij2005}^{2011} = P_{ij2011}^{2011} \times \frac{p_{ij2005} / p_{ij2011}}{p_{ib2005} / p_{ib2011}}$ for GDP and consumption.

2. Use the extrapolated PPPs for GDP to estimate total regional GDP, $Y_{2005}^{R2011} = \sum_{j \in R} \left(Y_{2005} / PPP_{Yj2005}^{2011} \right)$ for region *R* and analogously for consumption.

3. Allocate regional GDP to individual countries in proportion to the original ICP 2005 results: $\widehat{Y}_{j2005} = \frac{Y_{2005} / PPP_{Yj2005}^{2005}}{\sum_{j \in R} Y_{2005} / PPP_{Yj2005}^{2005}} \times Y_{2005}^{R2011}$.

4. The alternative GDP PPPs for 2005 are then computed as $P_{Yj2005}^{2005*} = Y_{j2005} / \widehat{Y}_{j2005}$.

---

[28] This difference between the GDP and consumption results could be due to the extrapolation using GDP deflators, rather than the more detailed extrapolation from Table 4.

[29] Table 7 shows that a regional breakdown of the global extrapolation results of Table 1 sketches a fairly comparable picture to the overall global extrapolation results. Note also that in particular the average difference and average squared difference for the Eurostat/OECD region is different from that reported in Table 5 even though the PPPs are the same. This is because the in Table 5, Germany is the numeraire for the Eurostat/OECD region, while the US is the numeraire in Table 7 and, as discussed earlier, the measures or not independent of the choice of numeraire.

This procedure is very similar to the Country Aggregation with Redistribution (CAR) approach used in ICP 2011 to link the regions. The fact that information from ICP 2011 is used in constructing these alternative 2005 PPPs makes it highly likely that using these alternative PPPs to predict 2011 PPPs will be an improvement over the original ICP 2005 PPPs. However, the results in Table 5 also illustrate that systematic differences could easily remain.

Table 6, ICP 2011 versus original and adjusted ICP 2005

	GDP		Consumption	
	Baseline	Adjusted ICP 2005	Baseline	Adjusted ICP 2005
Average difference	0.237***	-0.018	0.254***	-0.003
Average squared difference	0.113	0.026	0.122	0.017
Coefficient on log(expenditure/capita)	-0.020*	0.032***	-0.063***	0.019***
	(0.012)	(0.007)	(0.011)	(0.006)

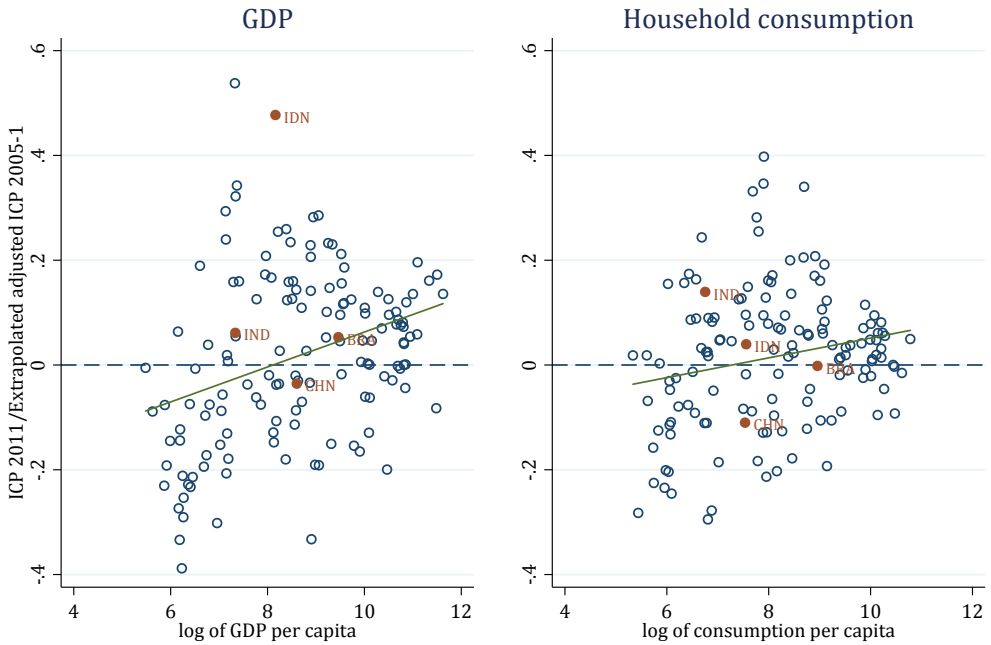
Notes: See notes of Table 1 for a general explanation of the statistics. Baseline figures correspond to the extrapolation of the original ICP 2005 from Table 1. The adjusted ICP 2005 is constructed following the procedure for relinking the ICP regions as described in the main text.

Table 6 shows that the adjusted ICP 2005 leads to drastically smaller differences relative to ICP 2011. When using the original ICP 2005 for extrapolation, the average country had an income or consumption per capita level that was about a quarter higher relative the US while using the adjusted ICP 2005 shows an average difference that does not differ significantly from zero. Similarly, the average squared difference is much reduced. There is still a systematic relationship between the difference and expenditure per capita and the sign has flipped: while compared with the original ICP 2005 the expenditure levels of lower-income countries were revised upwards by more under ICP 2011, the expenditure levels of lower-income countries are now revised downwards by more.

Figure 2 is the counterpart to Figure 1, but now based on the adjusted ICP 2005. This shows that, first, the differences are much more symmetric around zero, as already implied by Table 6. Furthermore, the figure suggests that the upward-

sloping relationship between the differences and (log) expenditure per capita are predominantly driven by countries at the lower end of the expenditure per capita. This is further confirmed by restricting the regression sample to countries with a GDP per capita level of at least $1100 (i.e. log(GDP/capita) of about 7): the regression coefficients based on the original ICP 2005 remain significant, while they are no longer significant for the adjusted ICP 2005. The adjusted ICP 2005 numbers are by no means now fully consistent with ICP 2011, but differences are less systematic than under the original ICP 2005. Furthermore, under the adjusted ICP 2005, 40-45 percent of countries have lower levels of expenditure per capita compared with 7-8 percent under the original ICP 2005. This means that the adjusted ICP 2005 is a less biased estimator for ICP 2011 and that there are fewer extreme differences that lead to the large average squared differences.

Figure 2, Differences between extrapolating the adjusted ICP 2005 and ICP 2011



Note: GDP and consumption per capita are in exchange-rate converted US$.
Extrapolated from adjusted ICP 2005 using GDP/consumption inflation rate of each country relative to the US.

To provide further perspective, Table 7 shows average differences and average squared differences for the original and adjusted ICP 2005 by ICP region.[30] The top row contains the same figures as in Table 6 and the subsequent rows shows

[30] The regression coefficients have been omitted for brevity but are available on request.

that the original ICP 2005 shows large average differences and average squared differences across the regions, with the exception of the Eurostat/OECD region. Especially for Western Asia, the 72-77 percent increase in ICP 2011 compared to the extrapolation from the original ICP 2005 is a shockingly large difference, but most other regions also show large differences. The adjusted ICP 2005 leads to much smaller differences – except in the Eurostat/OECD region and this is because the adjustment to ICP 2005 leaves the within-region PPPs unaffected. Even though the differences in Africa and Asia are still different from zero on average, the differences are of much less dramatic size at 5-10 compared with 20-40 percent under the original ICP 2005.

Table 7, ICP 2011 versus PPPs extrapolated from ICP 2005 – statistics by regions for the original and adjusted ICP 2005

	GDP				Consumption			
	$\overline{d}_Y$		$\overline{d}_Y^2$		$\overline{d}_C$		$\overline{d}_C^2$	
	Original	Adj.	Original	Adj.	Original	Adj.	Original	Adj.
Global	0.237***	-0.018	0.113	0.026	0.254***	-0.003	0.122	0.017
Africa	0.236***	-0.096***	0.118	0.042	0.333***	-0.049**	0.158	0.026
Asia	0.370***	0.095***	0.176	0.034	0.380***	0.052***	0.156	0.009
CIS	0.294***	-0.006	0.117	0.018	0.180**	0.020	0.081	0.037
Eurostat/ OECD	0.090***	0.090***	0.015	0.015	0.056***	0.056***	0.010	0.010
Latin America	0.169***	-0.009	0.037	0.006	0.160***	0.050	0.034	0.009
Western Asia	0.723***	0.045	0.559	0.015	0.769***	-0.030	0.612	0.007

Notes: "Original" refers to extrapolation from the original ICP 2005; "Adj." refers to extrapolation from the adjusted ICP 2005. the row 'Global' contains the summary statistics from Table 1 (original) and Table 6 (adjusted). The columns contain the average difference $\overline{d}_i$ and the average squared difference $\overline{d}_i^2$ for *i* equal to GDP, *Y*, or consumption, *C*. Extrapolation is done at the global level and each row presents the summary statistics for that region. See notes to Table 1 for further details and see World Bank (2014) for the country composition of each region.

Features of the adjusted ICP 2005

The adjusted version of ICP 2005 has two important advantages over the original:

1. From a methodological point of view, it is preferable to account for the price structure of all countries when linking the various regions rather than only the structure of a set of ring countries, and

2. Extrapolation of the adjusted ICP 2005 to 2011 provides price and income levels that are much closer to the actual ICP 2011 figures than the original ICP 2005 results.

These considerations could be enough to prefer the adjusted ICP 2005 over the original one, but before making such a decision, the consequences for features of cross-country income levels and income distribution should be considered. In this subsection I consider two indicators, namely the Balassa-Samuelson relationship and the extent of international income inequality.

Table 8 compares the linear Balassa-Samuelson relationship from equation (3) and the quadratic Balassa-Samuelson relationship for three sets of PPPs, namely the original ICP 2005, the adjusted ICP 2005 and ICP 2011. Results are shown for GDP PPPs and GDP per capita, but the results for consumption are very similar.[31] The linear Balassa-Samuelson relationship shows a considerably steeper slope for the adjusted ICP 2005 PPPs compared with the other two PPP sets. Indeed, while the Balassa-Samuelson relationship according to the original ICP 2005 results was not significantly different from the ICP 2011 relationship, the slope of the adjusted ICP 2005 relationship is significantly larger. However, for the quadratic relationship, the original and adjusted ICP 2005 are much more similar, while the ICP 2011 relationship stands out as comparatively different. As indicated earlier, the coefficients of these regressions have no clear relationship to any 'deep' or structural parameters, so any changes any differences cannot be used to confirm or reject a particular set of data. However, the comparative

[31] These are available on request.

similarity of results based on the adjusted ICP 2005 PPP indicates that those new PPPs do not show a notably different picture of relative prices across the world.

Table 8, Balassa-Samuelson relationship for ICP 2005, original and adjusted, and ICP 2011

	Original ICP 2005	Adjusted ICP 2005	ICP 2011
Linear			
log(GDP/capita)	0.215***	0.267***	0.203***
	(0.0124)	(0.0155)	(0.0157)
Constant	-2.368***	-2.979***	-2.299***
	(0.103)	(0.121)	(0.134)
R-squared	0.716	0.739	0.630
Quadratic			
log(GDP/capita)	-0.504***	-0.470***	-0.656***
	(0.112)	(0.131)	(0.156)
log(GDP/capita)	0.0440***	0.0451***	0.0501***
	(0.00686)	(0.00823)	(0.00936)
Constant	0.445	-0.0976	1.251*
	(0.446)	(0.507)	(0.640)
R-squared	0.788	0.789	0.719

Notes: dependent variable in the regression is the log of the GDP PPP over the exchange rate, explanatory variable is the log of (exchange-rate converted) GDP per capita; see also equation (3). The explanatory variable in the columns 'Original ICP 2005' and 'Adjusted ICP 2005' is 2005 log GDP/capita while the column 'ICP 2011' uses the 2011 values. Robust standard errors are in parentheses, each regression is run on 142 observations.*** p<0.01, ** p<0.05, * p<0.1.

But while the broad cross-country pattern of prices is not affected in a radical fashion, the adjusted ICP 2005 PPPs do imply a notably different pattern of cross-country inequality in GDP and consumption levels. Figure 3 shows Theil indexes of inequality based on the original ICP 2005 PPPs, the adjusted ICP 2005 PPPs and the ICP 2011 PPPs. In the terminology of e.g. Milanovic (2012) or Anand and Segal (2008), this is a so-called 'concept 2' measure of inequality in that it gives

larger weight to more populous countries but does not take into account that income and consumption are not distributed equally within countries. More formally, the Theil index is computed as:

$$(5) \qquad T = \sum_j s_j^y \left(\log\left(s_j^y\right) - \log\left(s_j^p\right) \right),$$

where s_j^y is the share of country *j* in global GDP and s_j^p is the share of country *j* in global population (the consumption Theil index is computed analogously). If each country's share of global GDP is equal to its share of global population, the Theil index goes to zero indicating perfect equality of (average) incomes.

Figure 3, Theil index of inequality for ICP 2005, original and adjusted, and ICP 2011

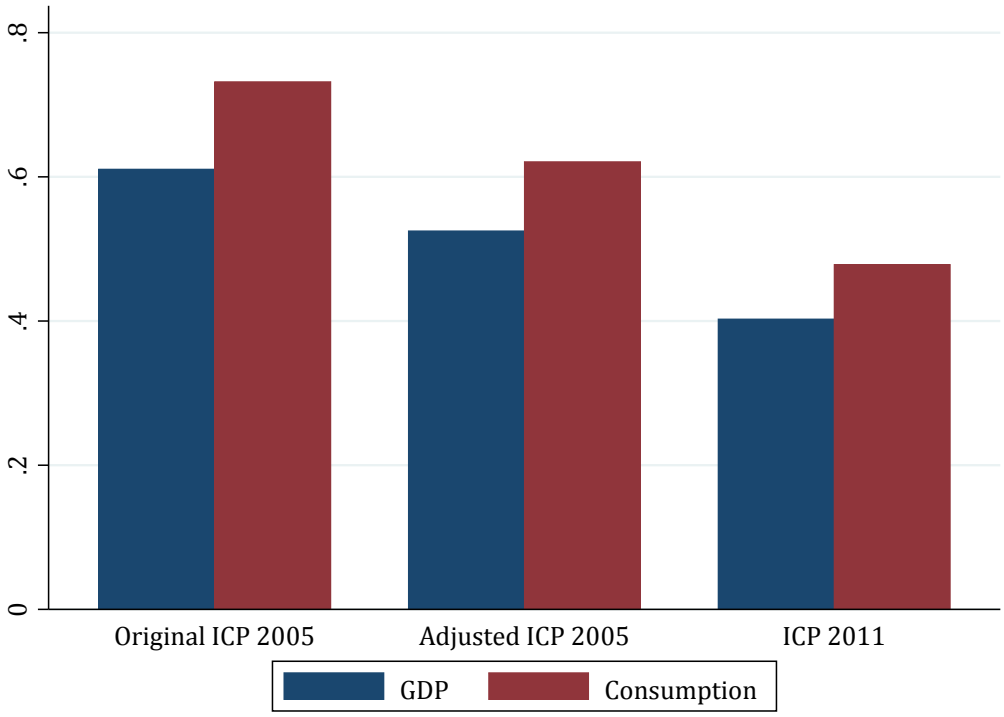


Figure 3 shows that the adjustment to ICP 2005 has a sizeable impact on income and consumption inequality. The GDP Theil index drops from 0.61 to 0.52 and the consumption Theil index shows a drop of similar size from 0.73 to 0.62. A downward shift for 2005 was to be expected as the original ICP 2005 implied large upwards revisions to income and consumption when moving to ICP 2011, while the adjusted ICP 2005 had an average revision of (close to) zero. But while

the downward shift in the Theil index is sizeable, ICP 2011 still implies a further decline in inequality by 0.12-0.15 points compared with the adjusted ICP 2005. According to the estimates of Milanovic (2012), inequality has been declining since ±2000. The adjusted ICP 2005 shifts some of that decline to the pre-2005 years compared with the 2005-2011 period. It would have been worrisome, though, if the inequality measures based on the adjusted ICP 2005 PPPs had been *lower* than based on ICP 2011. In that regard, the adjusted ICP 2005 PPPs fit the broader pattern as well as the original ICP 2005 PPPs.

A notable contributor to the decrease in inequality is the higher income level of China. In the adjusted ICP 2005, China's GDP level is 20 percent higher than in the original ICP 2005. This should be seen as a desirable feature as Feenstra, Ma, Neary and Rao (2013) have shown that ICP 2005 underestimated the size of China's economy by around 30 percent. The current adjustment moves China a notable step in that direction.

Concluding remarks

The publication of the ICP 2011 PPPs has prompted much debate on what the new view on world income levels implies for topics such as global poverty and what such revisions to earlier estimates imply for the future of ICP. If it is not possible to reach a broad level of consistency between different PPP benchmark years and relative inflation in between, it raises doubts about the usefulness of expending substantial resources on the estimation of PPPs.

This paper has attempted to close the gap between extrapolated PPPs based on ICP 2005 and ICP 2011. As a point of departure, I have taken ICP 2011 as the superior measurement exercise given that there have been clear methodological advances compared with ICP 2005, for instance in adjusting for the relevance and representativeness of individual products, in the treatment of hard-to-measure services such as rents and government services, and in how the regional comparisons are linked together to form a global comparison. I have shown that, given this faith in ICP 2011, it is not possible to form predictions or extrapolate based on the original ICP 2005 in such a way as to eliminate the inconsistency between the two benchmarks. This includes using the Balassa-Samuelson

relationship between prices and incomes as well extrapolation at a more detailed level than overall GDP or consumption.

My preferred alternative approach is to adjust the way in which regions were linked in ICP 2005 using information from ICP 2011. By using an approach that is similar to that applied in ICP 2011, most of the systematic differences between the two benchmark results can be overcome. The ex-post modification of ICP 2005 is certainly not a first-best solution, but the proposed modification leaves intact the most reliable element of ICP 2005, the comparison of prices within each region. The adjusted ICP 2005 shows a broadly similar relationship between price levels and income as the original ICP 2005 and ICP 2011 but a substantially reduced level of income inequality. Taking a broader view, this shifts some of the reductions in income inequality from the 2005-2011 period to the pre-2005 period.

In the broader context of the ICP, the lack of consistency between ICP 2005 and 2011 and, in my view, the need for an adjusted ICP 2005 to overcome this inconsistency is troubling. Yet it is also understandable. Although the first systematic PPP comparisons dates back to the late 1960s (Kravis, Heston and Summers, 1978) , ICP 2005 was the first PPP comparison with global scope and a design that could do justice to the wide variety of products consumed around the world. In that light, ICP 2011 represents a methodological evolution and holds out hope that a subsequent ICP comparison would require less effort to reconcile. A further helpful step would be to follow the lead of Eurostat and the OECD in having more frequent PPP comparisons. These would allow the maintenance of the existing statistical infrastructure, reduce the scope of one-time surprises from revisions and provide more scope for adjustment to changing insights on methodology.

References

Anand, Sudhir and Paul Segal (2008), "What do we know about global income inequality?" *Journal of Economic Literature* 46(1): 57-94.

Balassa, Bela (1964), "The Purchasing-Power Parity Doctrine: A Reappraisal" *Journal of Political Economy* 72: 584–96.

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

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

Dalgaard, Esben and Henrik Sørensen (2002), "Consistency between PPP Benchmarks and National Price and Volume Indices," Paper presented at the 27th General Conference of the International Association for Research on Income and Wealth, Sweden.

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

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

Deaton, Angus and Betina Aten (2014), "Trying to Understand the PPPs in ICP2011: Why are the Results so Different?" *NBER Working Papers*, no. 20244.

Diewert, W. Erwin (2009), "Similarity Indexes and Criteria for Spatial Linking" in D.S. Prasada Rao (ed.) *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, 183-216, Edward Elgar, Cheltenham, UK.

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.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2013), "The next generation of the Penn World Table" *NBER Working Paper* no. 19255.

Feenstra, Robert C., Hong Ma, J. Peter Neary and D.S. Prasada Rao (2013), "Who Shrunk China? Puzzles in the Measurement of Real GDP" *Economic Journal* 123(573): 1100-1129.

Harjargasht, Gholamreza and D.S. Prasada Rao (2006), "Systems of Index Numbers for International Price Comparisons Based on the Stochastic Approach" *CEPA Working Paper Series*, no. 04/2006.

Hill, Robert J. (2004), "Constructing Price Indexes across Space and Time: The Case of the European Union" *American Economic Review*, 94(5), 1379–410.

Inklaar, Robert (2013), "Price levels and economic growth: making sense of the revisions to data on real incomes: a comment" *Review of Income and Wealth* 59(4): 614-622.

Kravis, Irving, Alan Heston and Robert Summers (1978), *International Comparisons of Real Product Purchasing Power: United Nations International Comparison Project, Phase 2*, Johns Hopkins University Press, Baltimore.

Krijnse Locker, H. and H. D. Faerber (1984), "Space and Time Comparisons of Purchasing Power Parities and Real Values," Review of Income and Wealth, 30(2): 53–83.

McCarthy, Paul (2013), "Extrapolating PPPs and Comparing ICP Benchmark Results" (2013), in World Bank (ed.), *Measuring the Real Size of the World Economy*, Chapter 18, World Bank, Washington DC.

Milanovic, Branko (2012), "Global inequality recalculated and updated: the effect of new PPP estimates on global inequality and 2005 estimates" *Journal of Economic Inequality* 10(1): 1-18.

Rao, D. S. Prasada, Alicia Rambaldi, and Howard Doran (2010), "Extrapolation of Purchasing Power Parities using Multiple Benchmarks and Auxiliary

Information: A New Approach," *Review of Income and Wealth*, 56(1), S59–98.

Ravallion, Martin (2013), "Price levels and economic growth: making sense of the revisions to data on real incomes" *Review of Income and Wealth* 59(4): 593-613.

Ravallion, Martin (2014), "India's Puzzling New PPP", Center for Global Development, May 8, http://www.cgdev.org/blog/india's-puzzling-new-ppp.

Samuelson, Paul A. (1964), "Theoretical Notes on Trade Problems" *Review of Economics and Statistics* 46(2): 145–54.

Samuelson, Paul A. (1994), "Facets of Balassa-Samuelson Thirty Years Later" *Review of International Economics* 2(3): 201–26.

Summers, Robert and Alan Heston (1991), "The Penn World Table (Mark 5): An Extended Set of International Comparisons, 1950–1988" *Quarterly Journal of Economics* 106: 327–68.

Summers, Robert and Alan Heston (1993), "What Can Be Learned from Successive ICP Benchmark Estimates?" in Adam Szirmai and Dirk Pilat (eds), *Explaining Economic Growth; Essays in Honour of Angus Maddison*, North Holland, Amsterdam.

West, Kenneth D., "Forecast Evaluation," in Graham Elliot, Clive W. J. Granger, and Allan Timmermann (eds), *Handbook of Economic Forecasting, Volume 1*, 99–134, 2006.

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

World Bank (2014), *Purchasing Power Parities and Real Expenditure of World Economies – Summary of Results and Findings of the 2011 International Comparison Program*, World Bank, Washington DC.